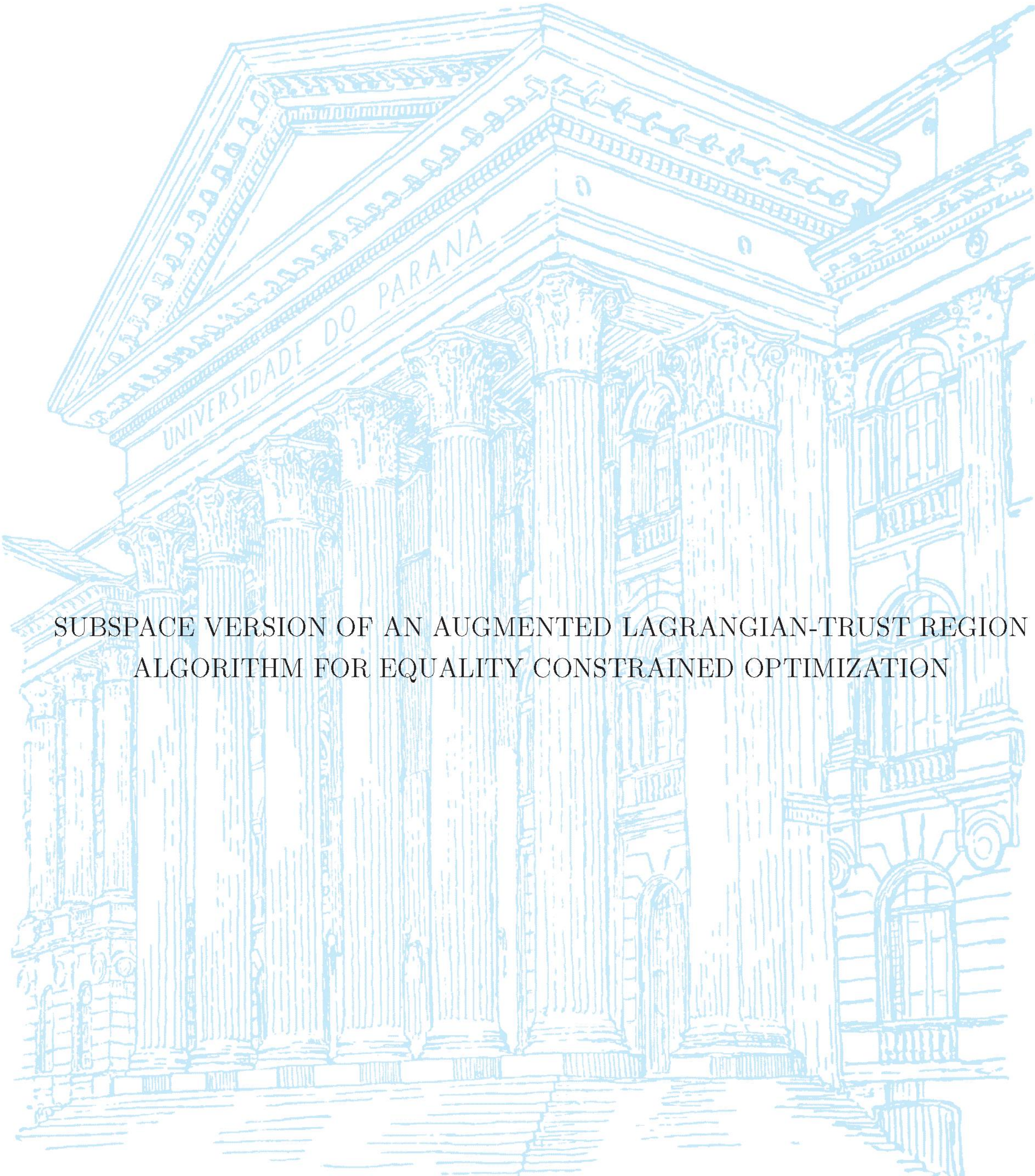
UNIVERSIDADE FEDERAL DO PARANÁ

CARINA MOREIRA COSTA




# SUBSPACE VERSION OF AN AUGMENTED LAGRANGIAN-TRUST REGION ALGORITHM FOR EQUALITY CONSTRAINED OPTIMIZATION

CURITIBA

2019

CARINA MOREIRA COSTA

# SUBSPACE VERSION OF AN AUGMENTED LAGRANGIAN-TRUST REGION ALGORITHM FOR EQUALITY CONSTRAINED OPTIMIZATION

Dissertação apresentada ao curso de Pós-Graduação em Matemática, Setor de Ciências Exatas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Geovani Nunes Grapiglia.

CURITIBA
2019

Catalogação na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

C837s Costa, Carina Moreira

Subspace version of an augmented lagrangian-trust region algorithm for equality constrained optimization [recurso eletrônico] / Carina Moreira Costa, 2019.

Dissertação (mestrado) - Programa de Pós-Graduação em Matemática, Setor de Ciências Exatas, Universidade Federal do Paraná
Orientador: Prof. Dr. Geovani Nunes Grapiglia.

1. Lagrange, Funções de. 2. Otimização matemática. 3. Matemática aplicada. I. Universidade Federal do Paraná. II. Grapiglia, Geovani Nunes. III. Título.

CDD 519.7

Bibliotecária: Vilma Machado CRB9/1563

MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MATEMÁTICA - 40001016041P1

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **CARINA MOREIRA COSTA** intitulada: **Subspace version of an Augmented Lagrangian-Trust Region Algorithm for Equality Constrained Optimization** , após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 28 de Fevereiro de 2019.

GEOVANI NUNES GRAPIGLIA
Presidente da Banca Examinadora (UFPR)

ROGER BEHLING
Avaliador Externo (UFSC)

CLOVIS CAESAR GONZAGA
Avaliador Externo (UFSC)

*Dedico esta dissertação aos meus pais, Marlene e Roque.*

## AGRADECIMENTOS

À Deus, pela proteção divina, pelo amor e pela força.

Aos meus pais, Marlene T. M. Costa e Roque G. M. Costa, pelos ensinamentos, pelo amor, pela compreensão, confiança e apoio para que eu sempre seguisse em frente.

Ao meu namorado, Ederson Machado, pelo incentivo, pela ajuda, pelo amor e compreensão.

À toda minha família, pelo carinho e incentivo.

Ao meu orientador, Professor Geovani Nunes Grapiglia, pelos valiosos ensinamentos, conselhos, motivação, confiança e pelas oportunidades.

Aos professores Jinyun Yuan, Rosangela Villwock, Sandro Marcos Guzzo e André Vicente, pelos ensinamentos e por me incentivarem a prosseguir com meus estudos.

À todos os meus professores, por ampliarem meus conhecimentos, pelos valiosos conselhos, pelo incentivo e por serem meus exemplos.

Aos professores, Clovis Caesar Gonzaga e Roger Behling, por aceitarem participar da banca examinadora e pelas contribuições que ajudaram muito a melhorar a versão final deste trabalho.

Aos amigos e colegas de estudo, pela ajuda, carinho e companheirismo em todos os momentos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

*"A mind that is stretched by a new experience can never go back to its old dimensions."*

Oliver Wendell Holmes Jr.

**RESUMO**

São apresentadas propriedades subespaciais para os subproblemas de região de confiança que aparecem no método Lagrangiano Aumentado-Região de Confiança proposto recentemente por Wang e Yuan (Optim. Methods Softw. 30, 559-582, 2015). Especificamente, quando as aproximações das Hessianas do Lagrangiano são atualizadas por fórmulas quase-Newton convenientemente escolhidas, mostra-se que o passo obtido do subproblema de região de confiança pertence ao subespaço gerado por todos os vetores gradientes da função objetivo e das restrições calculados até a iteração atual. Com base nesse resultado, propõe-se uma versão subespacial do método citado para problemas de otimização com restrições de igualdade de grande porte, nos quais o número de restrições é muito menor que o número de variáveis.



**Palavras-chave:** *Otimização com Restrições. Métodos de Lagrangiano Aumentado. Métodos de Região de Confiança. Métodos Subespacias.*

# ABSTRACT

Subspace properties are presented for the trust-region subproblems that appear in the Augmented Lagrangian-Trust-Region method recently proposed by Wang and Yuan (Optim. Methods Softw. 30, 559-582, 2015). Specifically, when the approximate Lagrangian Hessians are updated by suitable quasi-Newton formulas, it is shown that the trial step obtained from the trust-region subproblem belongs to the subspace spanned by all gradient vectors of the objective and of the constraints computed until the current iteration. Based on this result, a subspace version of the referred method is proposed for large-scale equality constrained optimization problems in which the number of constraints is much lower than the number of variables.



**Keywords:** *Constrained Optimization. Augmented Lagrangian Methods. Trust-Region Methods. Subspace Methods.*

## LIST OF FIGURES

## LIST OF TABLES

## NOTATION

| | |
|---|---|
| $\mathbb{R}$ | Set of real numbers |
| $\mathbb{R}^n$ | $n$-dimensional Euclidean space |
| $\mathbb{1}$ | Vector $[1, 1, \ldots, 1]^T \in \mathbb{R}^n$ |
| $\vert\,.\,\vert$ | Absolute value |
| $\Vert\,.\,\Vert$ | Vector or matrix norm |
| $\langle .\,, . \rangle$ | Inner Euclidean product |
| $S^\perp$ | Orthogonal complement of the set S with respect to inner Euclidean product |
| $\text{span}(S)$ | Vector subspace spanned by the set of vectors $S$ or by the columns of matrix $S$ |
| $P_\Omega(x)$ | Euclidean projection of point $x$ onto the closed and convex set $\Omega$ |
| $\text{Range}(A)$ | Range of matrix $A$ |
| $\text{Null}(A)$ | Null space of matrix $A$ |
| $\text{cond}(A)$ | Conditioning number of matrix $A$ |
| $A^\dagger$ | Pseudo-inverse of matrix $A$ |
| $\nabla f(x)$ | Gradient of the function $f$ at point $x$ |
| $\nabla^2 f(x)$ | Hessian matrix of the function $f$ at point $x$ |
| $J_f(x)$ | Jacobian matrix of the function $f$ at point $x$ |
| $B(x, \Delta)$ | Open ball with center at $x$ and radius $\Delta$, that is, $B(x, \Delta) = \{y \in \mathbb{R}^n; \Vert y - x \Vert < \Delta\}$ |
| $x_k$ | $k$th iterate |
| $\Delta_k$ | Radius of the trust region at iteration $k$ |
| $\lambda_k$ | Lagrange multiplier at iteration $k$ |
| $\sigma_k$ | Penalty parameter at iteration $k$ |
| $g_k$ | Gradient of the objective function at point $x_k$ |
| $B_k$ | Matrix in $\mathbb{R}^{n \times n}$ at point $x_k$ |
| $I_n$ | Identity matrix in $\mathbb{R}^{n \times n}$ |
| $q_k(.)$ | Model at iteration $k$ |

CONTENTS

# Chapter 1

# Introduction

Consider the equality constrained optimization problem

$$\text{minimize} \quad f(x), \tag{1.1}$$
$$\text{subject to} \quad c(x) = 0, \tag{1.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable. For convenience, the following notation will be used:

$$c(x) = [c_1(x), \ldots, c_m(x)]^T, \tag{1.3}$$
$$A(x) = J_c(x) \equiv [\nabla c_1(x) \ldots \nabla c_m(x)]^T, \tag{1.4}$$
$$g(x) = \nabla f(x). \tag{1.5}$$

Given $x_k \in \mathbb{R}^n$, we also use $f_k$ for $f(x_k)$, $c_k$ for $c(x_k)$, $A_k$ for $A(x_k)$ and $g_k$ for $g(x_k)$.

The Augmented Lagrangian-Trust Region method (ALTR) proposed in [13] is an iterative procedure to solve (1.1)-(1.2). At the beginning of the $k$th iteration, $x_k \in \mathbb{R}^n$, $\lambda_k \in \mathbb{R}^m$, $\sigma_k > 0$, $\Delta_k > 0$, $0 < \eta < \eta_1 < 1/2$ and $B_k \in \mathbb{R}^{n \times n}$ symmetric are available. If $x_k$ does not satisfy the KKT conditions, a trial step $s_k$ is computed by solving the trust-region subproblem

$$\min_{s \in \mathbb{R}^n} \quad q_k(s) \equiv g_k^T s - \lambda_k^T A_k s + \frac{1}{2} s^T B_k s + \frac{\sigma_k}{2} \|c_k + A_k s\|_2^2 \tag{1.6}$$
$$\text{s. t.} \quad \|s\|_2 \le \Delta_k. \tag{1.7}$$

Subproblem (1.6)-(1.7) is equivalent to

$$\min_{s \in \mathbb{R}^n} \quad \nabla_x L(x_k, \lambda_k)^T s + \frac{1}{2} s^T B_k s + \frac{\sigma_k}{2} \|c_k + A_k s\|_2^2 \tag{1.8}$$
$$\text{s. t.} \quad \|s\|_2 \le \Delta_k, \tag{1.9}$$

where $L(x, \lambda) \equiv f(x) - \lambda^T c(x)$ is the Lagrangian function associated to (1.1)-(1.2). Thus, if we denote the corresponding augmented Lagrangian function by

$$L(x,\lambda;\sigma) \equiv f(x) - \lambda^T c(x) + \frac{\sigma}{2}\|c(x)\|_2^2,$$

it follows that the objective in (1.8)-(1.9) is a quadratic approximation to

$$L(x_k+s,\lambda_k;\sigma_k) - L(x_k,\lambda_k;\sigma_k),$$

which is obtained as the sum of a quadratic approximation to $L(x_k+s,\lambda_k) - L(x_k,\lambda_k)$ with the squared norm of a linear approximation to $c(x_k+s)$. Matrix $B_k$ is an approximation to the Hessian $\nabla^2_{xx}L(x_k,\lambda_k)$. In summary, ALTR method computes $s_k$ by minimizing a quadratic approximation to the augmented Lagrangian, subject to a trust-region constraint. This combination of techniques justify the name given to this method.

The quality of $s_k$ is evaluated by using the ratio $\rho_k$ between the actual reduction of the augmented Lagrangian and the predicted reduction:

$$\rho_k = \frac{Ared_k}{Pred_k} \equiv \frac{L(x_k,\lambda_k;\sigma_k) - L(x_k+s_k,\lambda_k;\sigma_k)}{q_k(0) - q_k(s_k)}. \tag{1.10}$$

Similarly to traditional trust-region schemes, the next iterate is obtained as follows:

$$x_{k+1} = \begin{cases} x_k + s_k, & \text{if } \rho_k \geq \eta, \\ x_k, & \text{otherwise.} \end{cases}$$

Further, the trust-region radius $\Delta_{k+1}$ for the next iteration is given by the rule

$$\Delta_{k+1} = \begin{cases} \max\left\{\Delta_k,\ 1.5\|s_k\|_2\right\}, & \text{if } \rho_k \in [1-\eta_1, +\infty) \\ \Delta_k, & \text{if } \rho_k \in [\eta_1, 1-\eta_1) \\ \max\left\{0.5\Delta_k,\ 0.75\|s_k\|_2\right\}, & \text{if } \rho_k \in [\eta, \eta_1) \\ \|s_k\|_2/4, & \text{if } \rho_k < \eta. \end{cases}$$

In the ALTR method, the usual rules of augmented Lagrangian methods to update $\sigma_k$ and $\lambda_k$ were adapted. When $\sigma_k$ is very large, it forces $\|c(x_k+s_k)\|_2$ to be very small. Thus, the penalty parameter is updated taking into account the violation of the constraints. Specifically, the value of $\sigma_k$ is increased whenever

$$\nabla_x L(x_k,\lambda_k;\sigma_k) = 0 \text{ and } \|c_k\|_2 > 0.$$

Since $\sigma_k$ also influences the computation of $s_k$, the following inequality is checked:

$$Pred_k < \delta_k \sigma_k \min\left\{\Delta_k\|c_k\|_2, \|c_k\|_2^2\right\}, \tag{1.11}$$

where $\delta_k > 0$ is an auxiliary parameter. If (1.11) is satisfied, then $\sigma_k$ is increased in order to decrease the constraint violation. Regarding the vector $\lambda_k$ of Lagrange multipliers, it is updated only when $x_{k+1}$ is close to the feasible region, that is,

$$\|c_{k+1}\|_2 \leq R_k, \tag{1.12}$$

where $\{R_k\}$ is a nonincreasing sequence. If (1.12) is satisfied, $\lambda_k$ is updated in such a way that $\{\lambda_k\}$ is bounded. This is done by computing the auxiliary vector

$$\tilde{\lambda}_{k+1} = \arg\min_{\lambda \in \mathbb{R}^m} \|g_{k+1} - A_{k+1}^T \lambda\|_2^2. \tag{1.13}$$

After that, $\lambda_{k+1}$ is obtained by projecting $\tilde{\lambda}_{k+1}$ onto a compact box $[\lambda_{min}, \lambda_{max}]$:

$$\lambda_{k+1} = P_{[\lambda_{min}, \lambda_{max}]}(\tilde{\lambda}_{k+1}). \tag{1.14}$$

The ALTR method can be summarized as follows.

**Algorithm 1.1.** *Augmented Lagrangian Trust Region Method (ALTR)*

***Step 0*** *Given the constants $\beta \in (0,1)$, $\theta > 1$, $\lambda_{\min} < \lambda_{\max}$ and*

$$0 < \eta < \eta_1 < \frac{1}{2}, \quad R_1 = \max\left\{\|c(x_1)\|_2, 1\right\}.$$

*Given $x_1 \in \mathbb{R}^n$, $B_1 \in \mathbb{R}^{n \times n}$, $\lambda_1 \in \mathbb{R}^m$, $\sigma_1 > 1$, $\delta_1 > 0$ and $\Delta_1 > 0$, set $k :=$ 1.*

***Step 1*** *If $\|c_k\|_2 = 0$ and $P_{\mathrm{Null}(A_k)}(g_k) = 0$, Stop (return $x_k$ as a solution).*

***Step 2*** *Set $\sigma_k^1 := \sigma_k$ and $j := 1$.*

*While $\nabla_x L(x_k, \lambda_k; \sigma_k^j) = 0$ and $\|c_k\|_2 > 0$, set*

$$\sigma_k^{j+1} := \theta \sigma_k^j \quad \textit{and} \quad j := j + 1 \tag{1.15}$$

*End (While). Set $\sigma_k = \sigma_k^j$. Compute $s_k$ by solving (1.6)-(1.7).*

***Step 3*** *Compute the ratio $\rho_k$ defined in (1.10). If $\rho_k \geq \eta$, go to Step 4, otherwise, set $\Delta_{k+1} = \frac{\|s_k\|_2}{4}$, $x_{k+1} = x_k$, $k := k + 1$ and go back to Step 2.*

***Step 4*** *If (1.11) is satisfied, then set*

$$\sigma_{k+1} = 2\sigma_k, \ \textit{and}\ \delta_{k+1} = \delta_k/4. \tag{1.16}$$

*Otherwise, set $\sigma_{k+1} = \sigma_k$ and $\delta_{k+1} = \delta_k$.*

*If $\|c_{k+1}\|_2 \leq R_k$, then compute $\lambda_{k+1}$ by (1.13)-(1.14) and set $R_{k+1} = \beta R_k$.*

*Otherwise, set $\lambda_{k+1} = \lambda_k$ and $R_{k+1} = R_k$.*

***Step 5*** *Set $x_{k+1} = x_k + s_k$ and*

$$\Delta_{k+1} = \begin{cases} \max\{\Delta_k, 1.5\,\|s_k\|_2\}, & \textit{if } \rho_k \in [1-\eta_1, +\infty), \\ \Delta_k, & \textit{if } \rho_k \in [\eta_1, 1-\eta_1), \\ \max\{0.5\Delta_k, 0.75\,\|s_k\|_2\}, & \textit{if } \rho_k \in [\eta, \eta_1). \end{cases}$$

*Compute $f_{k+1}$, $g_{k+1}$, $c_{k+1}$ and $A_{k+1}$, generate $B_{k+1}$, set $k := k+1$ and go back to Step 1.*

By developing the squared norm in $q_k(s)$, we see that

$$q_k(s) = (g_k - A_k^T\lambda_k + \sigma_k A_k^T c_k)^T s + \frac{1}{2}s^T(B_k + \sigma_k A_k^T A_k)s + \frac{\sigma_k}{2}\|c_k\|_2^2.$$

Then, denoting

$$\hat{g}_k = g_k - A_k^T\lambda_k + \sigma_k A_k^T c_k$$

and

$$\hat{B}_k = B_k + \sigma_k A_k^T A_k,$$

it follows that (1.6)-(1.7) is equivalent to the problem

$$\min_{s\in\mathbb{R}^n} \quad q_k(s) \equiv (\hat{g}_k)^T s + \frac{1}{2}s^T\hat{B}_k s \tag{1.17}$$

$$\text{s. t.} \quad \|s\|_2 \leq \Delta_k, \tag{1.18}$$

which is a standard trust-region subproblem. The efficiency of ALTR strongly depends on how accurately the trust-region subproblems are solved. If one is willing to solve (1.17)-(1.18) nearly exactly, the usual procedure is the iterative method of Moré and Sorensen [7]. However, at each iteration, this method requires at least one Cholesky factorization of matrices of the form $\hat{B}_k + \mu I_n$. This makes its use computationally expensive and prohibitive for large-scale problems (i.e., problems with $n$ very large).

Motivated by the subspace trust-region methods proposed by Wang and Yuan [12] and by Grapiglia, Yuan and Yuan [4], in this work we explore subspace properties of the trust-region subproblem (1.6)-(1.7) when the matrices $B_k$ are updated by quasi-Newton formulas. By adapting the analysis presented in [12, 4], it is found that any solution $s_k$ of (1.6)-(1.7) belongs to the subspace

$$G_k = \text{span}\left(\cup_{i=1}^{k}\{\nabla c_1(x_i), \ldots, \nabla c_m(x_i), g_i\}\right).$$

Therefore, we can restrict the search for $s_k$ to $G_k$. Note that $\dim(G_k) \leq k(m+1)$. Thus, if problem (1.1)-(1.2) has a small number of constraints ($m << n$), then the use of subspace $G_k$ in early iterations ($k << n$) may result in a significant reduction of the computational cost to solve the corresponding subproblem. Based on this observation, we propose a subspace version of ALTR for large-scale equality constrained problems in which the number of constraints is much lower than the number of variables.

This dissertation is organized as follows. In Chapter 2, the equivalence between the *full-space* trust-region subproblem and its subspace counterpart is proved, and the corresponding subspace version of ALTR is presented. The global convergence analysis of the subspace ALTR is given in Chapter 3. Finally, preliminary numerical results are reported in Chapter 4.

# Chapter 2

# Subspace Version of ALTR

In this chapter, we propose a subspace version of the ALTR method. First, we study subspace properties of the trust-region subproblem (1.6)-(1.7). All the results are obtained by adapting the analysis from [12, 4].

## 2.1 Subspace Properties

The following lemma characterizes the global solutions of the subproblem (1.6)-(1.7).

**Lemma 2.1.** *A vector $s_k \in \mathbb{R}^n$ is a global solution of (1.6)-(1.7) if, and only if, there exists a scalar $\mu_k \geq 0$ such that*

$$\begin{aligned}(\hat{B}_k + \mu_k I_n)s_k &= -\hat{g}_k,\\ \mu_k(\Delta_k - \|s_k\|_2) &= 0,\\ \|s_k\|_2 &\leq \Delta_k,\end{aligned}$$

*and $(\hat{B}_k + \mu_k I_n)$ is positive semidefinite.*

*Proof:* Due to the equivalence between (1.6)-(1.7) and (1.17)-(1.18), the result follows from Theorem 6.1.2 in [11]. ■

The next lemma establishes sufficient conditions under which any global solution to (1.6)-(1.7) belongs to a subspace $S_k$. Its proof is an adaptation of the proof of Lemma 2.2 in [4].

**Lemma 2.2.** *Let $S_k$ a subspace $r$-dimensional $(1 \leq r \leq n)$ of $\mathbb{R}^n$ and $Z_k \in \mathbb{R}^{n\times r}$ a matrix whose columns form an orthonormal basis of the subspace $S_k$, namely,*

$$S_k = \operatorname{span}(Z_k) \quad \textit{and} \quad Z_k^T Z_k = I_r. \tag{2.1}$$

*Suppose that*

$$\{\nabla c_1(x_k), \ldots, \nabla c_m(x_k), g_k\} \subset S_k, \tag{2.2}$$

*and* $B_k \in \mathbb{R}^{n\times n}$ *is a symmetric matrix satisfying*

$$B_k u = \alpha u, \ \forall u \in S_k^{\perp}, \tag{2.3}$$

*where* $\alpha > 0$. *Then, the vector* $s_k \in \mathbb{R}^n$ *is a solution of (1.6)-(1.7) if, and only if,* $s_k = Z_k \bar{s}_k \in S_k$ *where* $\bar{s}_k \in \mathbb{R}^r$ *is a solution to the following problem:*

$$\min_{\bar{s}\in\mathbb{R}^r} \ \bar{q}_k(\bar{s}) \equiv \bar{g}_k^T \bar{s} - \lambda_k^T \bar{A}_k \bar{s} + \frac{1}{2}\bar{s}^T \bar{B}_k \bar{s} + \frac{\sigma_k}{2}\left\|c_k + \bar{A}_k \bar{s}\right\|_2^2 \tag{2.4}$$

$$s.\ t. \ \ \|\bar{s}\|_2 \leq \Delta_k, \tag{2.5}$$

*where* $\bar{g}_k = Z_k^T g_k$, $\bar{A}_k = A_k Z_k$ *and* $\bar{B}_k = Z_k^T B_k Z_k$.

*Proof:* Let $U_k \in \mathbb{R}^{n\times(n-r)}$ be a matrix whose columns are an orthonormal basis of subspace $S_k^{\perp}$. Then, $[Z_k \ U_k] \in \mathbb{R}^{n\times n}$ is an orthogonal matrix whose columns are an orthonormal basis of $\mathbb{R}^n$. Consequently, for each $s \in \mathbb{R}^n$, there exists an unique pair of vectors $(\bar{s}, u) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$ such that $s = Z_k\bar{s} + U_k u$. From (1.6) it follows that

$$\begin{aligned}
q_k(s) &= q_k(Z_k\bar{s} + U_k u) \\
&= g_k^T(Z_k\bar{s} + U_k u) - \lambda_k^T A_k(Z_k\bar{s} + U_k u) + \frac{1}{2}(Z_k\bar{s} + U_k u)^T B_k (Z_k\bar{s} + U_k u) \\
&\quad + \frac{\sigma_k}{2}\|c_k + A_k(Z_k\bar{s} + U_k u)\|_2^2 \\
&= (Z_k^T g_k)^T \bar{s} + g_k^T U_k u - \lambda_k^T A_k Z_k \bar{s} - \lambda_k^T A_k U_k u + \frac{1}{2}\bar{s}^T Z_k^T B_k Z_k \bar{s} + \frac{1}{2}\bar{s}^T Z_k^T B_k U_k u \\
&\quad + \frac{1}{2}u^T U_k^T B_k Z_k \bar{s} + \frac{1}{2}u^T U_k^T B_k U_k u + \frac{\sigma_k}{2}\|c_k + A_k Z_k \bar{s} + A_k U_k u\|_2^2 \\
&= \bar{g}_k^T \bar{s} + g_k^T U_k u - \lambda_k^T \bar{A}_k \bar{s} - \lambda_k^T A_k U_k u + \frac{1}{2}\bar{s}^T \bar{B}_k \bar{s} + \frac{1}{2}\bar{s}^T Z_k^T B_k U_k u \\
&\quad + \frac{1}{2}u^T U_k^T B_k Z_k \bar{s} + \frac{1}{2}u^T U_k^T B_k U_k u + \frac{\sigma_k}{2}\left\|c_k + \bar{A}_k \bar{s} + A_k U_k u\right\|_2^2,
\end{aligned} \tag{2.6}$$

where $\bar{g}_k = Z_k^T g_k$, $\bar{A}_k = A_k Z_k$ and $\bar{B}_k = Z_k^T B_k Z_k$. Since $g_k \in S_k$ and the columns of $U_k$ are vectors in $S_k^{\perp}$, we obtain

$$g_k^T U_k = 0, \ Z_k^T B_k U_k = \alpha Z_k^T U_k = 0, \ U_k^T B_k Z_k = \alpha U_k^T Z_k = 0 \text{ and } U_k^T B_k U_k = \alpha I_{n-r}, \tag{2.7}$$

where we also used the hypothesis that $B_k$ is symmetric. From the fact that the rows of $A_k$ are the vectors $\nabla c_i(x_k) \in S_k$, and the columns of $U_k$ are vectors in $S_k^{\perp}$, it follows that

$$A_k U_k = 0. \tag{2.8}$$

Hence, (2.6)-(2.8) imply that

$$q_k(s) = q_k(\bar{s}, u) = \left( \bar{g}_k^T \bar{s} - \lambda_k^T \bar{A}_k \bar{s} + \frac{1}{2} \bar{s}^T \bar{B}_k \bar{s} + \frac{\sigma_k}{2} \left\| c_k + \bar{A}_k \bar{s} \right\|_2^2 \right) + \frac{1}{2} \alpha \left\| u \right\|_2^2 . \tag{2.9}$$

In addition, since $Z_k^T U_k = 0$, we have

$$\|s\|_2^2 = \|Z_k \bar{s} + U_k u\|_2^2 = \|\bar{s}\|_2^2 + \|u\|_2^2 . \tag{2.10}$$

Therefore, (2.9) and (2.10) imply that solving the subproblem (1.6)-(1.7) is equivalent to solve the subproblem

$$\min_{(\bar{s},u) \in \mathbb{R}^r \times \mathbb{R}^{n-r}} \quad \left( \bar{g}_k^T \bar{s} - \lambda_k^T \bar{A}_k \bar{s} + \tfrac{1}{2} \bar{s}^T \bar{B}_k \bar{s} + \tfrac{\sigma_k}{2} \left\| c_k + \bar{A}_k \bar{s} \right\|_2^2 \right) + \tfrac{1}{2} \alpha \left\| u \right\|_2^2 \tag{2.11}$$

$$\text{s. t.} \quad \|\bar{s}\|_2^2 + \|u\|_2^2 \leq \Delta_k^2, \tag{2.12}$$

with respect to $s = Z_k \bar{s} + U_k u$.

Let $s_k = Z_k \bar{s}_k + U_k u_k$ a solution of (1.6)-(1.7). We will show that $u_k = 0$. For that, suppose by contradiction that $u_k \neq 0 \in \mathbb{R}^{n-r}$. Since $s_k$ is an optimal solution, we have $\|s_k\|_2^2 \leq \Delta_k^2$ and

$$q_k(s_k) \leq q_k(s)$$

for all $s \in \mathbb{R}^n$ satisfying $\|s\|_2^2 \leq \Delta_k^2$. In particular,

$$q_k(s_k) \leq q_k(s_k^*), \tag{2.13}$$

where $s_k^* = Z_k \bar{s}_k$. However, since $u_k \neq 0 \in \mathbb{R}^{n-r}$ and $\alpha > 0$, from (2.9) it follows that

$$\begin{aligned}
q_k(s_k) &= q_k(\bar{s}_k, u_k) \\
&= \bar{g}_k^T \bar{s}_k - \lambda_k^T \bar{A}_k \bar{s}_k + \frac{1}{2} \bar{s}_k^T \bar{B}_k \bar{s}_k + \frac{\sigma_k}{2} \left\| c_k + \bar{A}_k \bar{s}_k \right\|_2^2 + \frac{1}{2} \alpha \left\| u_k \right\|_2^2 \\
&> \bar{g}_k^T \bar{s}_k - \lambda_k^T \bar{A}_k \bar{s}_k + \frac{1}{2} \bar{s}_k^T \bar{B}_k \bar{s}_k + \frac{\sigma_k}{2} \left\| c_k + \bar{A}_k \bar{s}_k \right\|_2^2 \\
&= q_k(s_k^*),
\end{aligned}$$

which contradicts (2.13). Therefore, we must have $u_k = 0 \in \mathbb{R}^{n-r}$. This shows that if $s_k$ is a solution of (1.6)-(1.7) then $s_k = Z_k \bar{s}_k \in S_k$. The fact that $\bar{s}_k$ is a solution of (2.4)-(2.5) follows from the equivalence between (1.6)-(1.7) and (2.11)-(2.12) with $u = 0$.

Reciprocally, if $\bar{s}_k$ is a solution of (2.4)-(2.5) then $(\bar{s}_k, 0) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$ is a solution of (2.11)-(2.12) and, consequently, $s_k = Z_k \bar{s}_k$ is a solution of (1.6)-(1.7). ■

**Remark 2.1.** *By Lemma 2.2, if assumptions (2.2)-(2.3) are satisfied, then we can solve subproblem (2.4)-(2.5) in $\mathbb{R}^r$ instead of solving subproblem (1.6)-(1.7) in $\mathbb{R}^n$. When $r << n$, this can give a significant reduction in the computational cost to obtain $s_k$.*

The next lemma provides a subspace that satisfies the assumptions of Lemma 2.2. Its proof is an adaptation of the proof of Lemma 2.3 in [4].

**Lemma 2.3.** *Suppose that $B_1 = \alpha I_n$, with $\alpha > 0$, and that $B_k$ is the $k$th update matrix given by a quasi-Newton formula chosen from the PSB and the Broyden family. Let $s_k \in \mathbb{R}^n$ be a solution of (1.6)-(1.7) and*

$$G_k = \operatorname{span}\left(\bigcup_{i=1}^{k}\{\nabla c_1(x_i), \dots, \nabla c_m(x_i), g_i\}\right). \tag{2.14}$$

*Then, for all $k$, we have $s_k \in G_k$ and $B_k u = \alpha u$ for all $u \in G_k^{\perp}$.*

*Proof:* The PSB formula and the Broyden family are specified, respectively, by

$$\begin{aligned} B_{k+1}^{(PSB)} = B_k^{(PSB)} &+ \frac{(y_k - B_k^{(PSB)} s_k)s_k^T + s_k(y_k - B_k^{(PSB)} s_k)^T}{s_k^T s_k} \\ &- \frac{(y_k - B_k^{(PSB)} s_k)^T s_k s_k s_k^T}{(s_k^T s_k)^2} \end{aligned} \tag{2.15}$$

$$B_{k+1}^{(B)} = B_k^{(B)} - \frac{B_k^{(B)} s_k s_k^T B_k^{(B)}}{s_k^T B_k^{(B)} s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \theta_k (s_k^T B_k^{(B)} s_k) w_k w_k^T \tag{2.16}$$

where $s_k = x_{k+1} - x_k$, $y_k = (g_{k+1} - g_k) - (A_{k+1} - A_k)^T \lambda_k$ and

$$w_k = \frac{y_k}{s_k^T y_k} - \frac{B_k^{(B)} s_k}{s_k^T B_k^{(B)} s_k}.$$

We will prove the result by induction over $k$. For $k = 1$ we have $B_1 = \alpha I_n$, $\alpha > 0$, $G_1 = \operatorname{span}\{\nabla c_1(x_1), \dots, \nabla c_m(x_1), g_1\}$ and $A_1 = [\nabla c_1(x_1) \dots \nabla c_m(x_1)]^T$. Let $s_1 \in \mathbb{R}^n$ be solution of (1.6)-(1.7) with $k = 1$. Then, by Lemma 2.1, there exists $\mu_1 \geq 0$ such that

$$\begin{aligned} & (\hat{B}_1 + \mu_1 I_n)s_1 = -\hat{g}_1 \\ \Longrightarrow\ & (B_1 + \sigma_1 A_1^T A_1 + \mu_1 I_n)s_1 = -(g_1 - A_1^T \lambda_1 + \sigma_1 A_1^T c_1) \\ \Longrightarrow\ & (\alpha I_n + \mu_1 I_n)s_1 = -(g_1 - A_1^T \lambda_1 + \sigma_1 A_1^T c_1 + \sigma_1 A_1^T A_1 s_1) \\ \Longrightarrow\ & s_1 = -(\alpha + \mu_1)^{-1}(g_1 - A_1^T \lambda_1 + \sigma_1 A_1^T c_1 + \sigma_1 A_1^T A_1 s_1) \\ \Longrightarrow\ & s_1 \in G_1, \end{aligned}$$

where the last line is true because $g_1$, $A_1^T \lambda_1$, $A_1^T c_1$ and $A_1^T A_1 s_1 \in G_1$. Moreover,

$$B_1^{(PSB)} u = B_1^{(B)} u = (\alpha I_n) u = \alpha u, \quad \forall u \in G_1^{\perp}.$$

Hence, the lemma is true for $k=1$. Assume that the lemma is true for $k=i$, that is,

$$s_i \in G_i, \tag{2.17}$$

and

$$B_i^{(PSB)} u = B_i^{(B)} u = \alpha u, \quad \forall u \in G_i^{\perp}. \tag{2.18}$$

Consider $\tilde{u} \in G_{i+1}^{\perp}$. Then, we also have $\tilde{u} \in G_i^{\perp}$ (since $G_i \subset G_{i+1} \Rightarrow G_{i+1}^{\perp} \subset G_i^{\perp}$). Note that $y_i \in G_{i+1}$ and matrices $B_i^{(PSB)}$ and $B_i^{(B)}$ are symmetric. Thus, it follows from (2.17) and (2.18) that

$$\begin{aligned}
B_{i+1}^{(PSB)} \tilde{u} &= B_i^{(PSB)} \tilde{u} + \frac{\left((y_i - B_i^{(PSB)} s_i) s_i^T + s_i (y_i - B_i^{(PSB)} s_i)^T\right) \tilde{u}}{s_i^T s_i} \\
&\quad - \frac{(y_i - B_i^{(PSB)} s_i)^T s_i s_i s_i^T \tilde{u}}{(s_i^T s_i)^2} \\
&= \alpha \tilde{u} + \frac{(y_i - B_i^{(PSB)} s_i) s_i^T \tilde{u} + s_i (y_i^T \tilde{u} - s_i^T B_i^{(PSB)} \tilde{u})}{s_i^T s_i} \\
&= \alpha \tilde{u} - \alpha \frac{s_i s_i^T \tilde{u}}{s_i^T s_i} \\
&= \alpha \tilde{u},
\end{aligned} \tag{2.19}$$

and

$$\begin{aligned}
B_{i+1}^{(B)} \tilde{u} &= B_i^{(B)} \tilde{u} - \frac{B_i^{(B)} s_i s_i^T B_i^{(B)} \tilde{u}}{s_i^T B_i^{(B)} s_i} + \frac{y_i y_i^T \tilde{u}}{s_i^T y_i} + \theta_i (s_i^T B_i^{(B)} s_i) w_i w_i^T \tilde{u} \\
&= \alpha \tilde{u} - \frac{\alpha B_i^{(B)} s_i s_i^T \tilde{u}}{s_i^T B_i^{(B)} s_i} + \theta_i (s_i^T B_i^{(B)} s_i) w_i \left( \frac{y_i^T}{s_i^T y_i} - \frac{s_i^T B_i^{(B)}}{s_i^T B_i^{(B)} s_i} \right) \tilde{u} \\
&= \alpha \tilde{u} + \theta_i (s_i^T B_i^{(B)} s_i) w_i \left( \frac{y_i^T \tilde{u}}{s_i^T y_i} - \frac{s_i^T B_i^{(B)} \tilde{u}}{s_i^T B_i^{(B)} s_i} \right) \\
&= \alpha \tilde{u} - \alpha \theta_i (s_i^T B_i^{(B)} s_i) w_i \frac{s_i^T \tilde{u}}{s_i^T B_i^{(B)} s_i} \\
&= \alpha \tilde{u}.
\end{aligned}$$

Since $\tilde{u} \in G_{i+1}^{\perp}$ is arbitrary, this proves that

$$B_{i+1}^{(PSB)} u = B_{i+1}^{(B)} u = \alpha u, \quad \forall u \in G_{i+1}^{\perp}. \tag{2.20}$$

Now, let $s_{i+1}$ be a solution of the subproblem (1.6)-(1.7) for $k=i+1$. Then, by

$$\{\nabla c_1(x_{i+1}), \cdots, \nabla c_m(x_{i+1}), g_{i+1}\} \subset G_{i+1},$$

equation (2.20) and Lemma 2.2 (with $k = i+1$) we conclude that $s_{i+1} = Z_{i+1}\bar{s}_{i+1} \in G_{i+1}$ (where $\bar{s}_{i+1}$ is a solution of the subproblem (2.4)-(2.5) for $k = i+1$, and $Z_{i+1}$ is a matrix whose columns are an orthonormal basis of $G_{i+1}$). The proof is complete. ∎

**Remark 2.2.** *For the further analysis, it is useful to see that*

$$B_k u = \alpha u, \quad \forall u \in G_k^{\perp} \quad \Rightarrow \quad B_k z \in G_k, \quad \forall z \in G_k. \tag{2.21}$$

*Indeed, given $z \in G_k$ and $u \in G_k^{\perp}$ arbitrary, as $B_k$ is a symmetric matrix, we have*

$$(B_k z)^T u = z^T B_k^T u = z^T B_k u = \alpha z^T u = 0.$$

*Thus, $B_k z \in (G_k^{\perp})^{\perp} = G_k$, for all $z \in G_k$.*

Now, combining Lemmas 2.2 and 2.3, we obtain the following theorem.

**Theorem 2.4.** *Let $Z_k$ be a matrix whose columns are an orthonormal basis of the subspace $G_k$ given in (2.14). Suppose that $B_1 = \alpha I_n$ $(\alpha > 0)$ and that $B_k$ is the $k$th updated matrix given by a quasi-Newton formula chosen from the PSB and the Broyden family. Let $s_k$ be a solution of the subproblem (1.6)-(1.7). Then, there exists a solution $\bar{s}_k$ of (2.4)-(2.5) such that $s_k = Z_k\bar{s}_k$, which implies $s_k \in G_k$. Reciprocally, if $\bar{s}_k$ is a solution of (2.4)-(2.5), then $s_k = Z_k\bar{s}_k$ is a solution of (1.6)-(1.7).*

The following lemma establishes that the approximate Hessian matrix $B_k$ can be updated in the subspace $G_k$. Its proof is due to Siegel [10], Gill and Leonard [2]. We give it here for completeness.

**Lemma 2.5.** *Let $Z \in \mathbb{R}^{n\times r}$ be a matrix whose columns are orthonormal. Suppose that $s_k \in \operatorname{span}(Z)$, and that matrix $B_{k+1} = Update\,(B_k, s_k, y_k)$ is obtained by the PSB formula or by some formula from the Broyden family. Then, denoting $\bar{B}_{k+1} = Z^T B_{k+1} Z$, $\tilde{B}_k = Z^T B_k Z$, $\tilde{s}_k = Z^T s_k$ and $\tilde{y}_k = Z^T y_k$, we have $\bar{B}_{k+1} = Update\left(\tilde{B}_k, \tilde{s}_k, \tilde{y}_k\right)$.*

*Proof:* Note that $s_k \in \operatorname{span}(Z) \implies s_k = ZZ^T s_k$. Then,

$$\begin{aligned}
s_k^T y_k &= \left(ZZ^T s_k\right)^T y_k = \left(Z^T s_k\right)^T Z^T y_k = \tilde{s}_k^T \tilde{y}_k \\
s_k^T B_k s_k &= \left(ZZ^T s_k\right)^T B_k \left(ZZ^T s_k\right) = \left(Z^T s_k\right)^T Z^T B_k Z \left(Z^T s_k\right) = \tilde{s}_k^T \tilde{B}_k \tilde{s}_k \\
Z^T B_k s_k &= Z^T B_k Z \left(Z^T s_k\right) = \tilde{B}_k \tilde{s}_k.
\end{aligned}$$

Then, the conclusion follows by multiplying (2.15) and (2.16) by $Z^T$ for the left and by $Z$ for the right. ∎

**Remark 2.3.** *By Theorem 2.4, we can solve the subproblem (1.6)-(1.7) by solving (2.4)-(2.5) in the subspace $G_k$, provided that $B_1 = \alpha I_n$ and a suitable quasi-Newton formula* *is used to update $B_k$. Moreover, it follows from Lemma 2.5 that the reduced matrix $\bar{B}_k = Z_k^T B_k Z_k$ of $B_k$ in the subspace $G_k$ can be obtained by updating the matrix $\tilde{B}_{k-1} = Z_k^T B_{k-1} Z_k$, where $Z_k$ is the matrix whose columns are an orthonormal basis of the subspace $G_k$. These subspace properties can be explored to reduce the computational cost to compute $s_k$ when $m << n$ and the dimension of the subspace $G_k$ remains much smaller than $n$.*

## 2.2 Subspace Algorithm

In this section, we present a subspace version of ALTR based on the subspace properties described in the previous section. Suppose that at the $k$th iteration, $Z_k \in \mathbb{R}^{n \times r_k}$ has been obtained, which is an orthonormal basis matrix of the subspace $G_k$. Further, suppose that $\bar{s}_k$ is obtained by solving the subproblem (2.4)-(2.5) and $s_k = Z_k \bar{s}_k$, $x_{k+1} = x_k + s_k$, $g_{k+1} = \nabla f(x_{k+1})$ and $A_{k+1} = [\nabla c_1(x_{k+1}) \ldots \nabla c_m(x_{k+1})]^T$. Then, we have to compute $Z_{k+1}$, $\bar{g}_{k+1} = Z_{k+1}^T g_{k+1}$, $\bar{A}_{k+1} = A_{k+1} Z_{k+1}$ and $\bar{B}_{k+1} = Z_{k+1}^T B_{k+1} Z_{k+1}$ for the next iteration. To obtain $Z_{k+1}$, we use the procedure of Gram-Schmidt with reorthogonalization (see Section 2 in Daniel et al. [1]). For this purpose, consider the notation:

$$p_j^{(k+1)} = \begin{cases} \nabla c_j(x_{k+1}), & \text{if } j = 1, ..., m \\ g_{k+1}, & \text{if } j = m+1. \end{cases} \tag{2.22}$$

Let $W_1 = Z_k$ and $t_1 = r_k$, where $r_k$ denotes the number of columns of $Z_k$. For $j = 1, ..., m+1$, by the reorthogonalization procedure, compute the decomposition

$$p_j^{(k+1)} = W_j u_j^{(k)} + \tau_j^{(k+1)} z_j^{(k+1)} \tag{2.23}$$

where

$$u_j^{(k)} = W_j^T p_j^{(k+1)}, \quad z_j^{(k+1)} \perp \operatorname{span}(W_j), \quad \left\| z_j^{(k+1)} \right\|_2 = 1, \tag{2.24}$$

and

$$\tau_j^{(k+1)} = \left\| p_j^{(k+1)} - W_j u_j^{(k)} \right\|_2 = \left\| (I - W_j W_j^T) p_j^{(k+1)} \right\|_2 \geq 0. \tag{2.25}$$

If $\tau_j^{(k+1)} > 0$, it follows that $p_j^{(k+1)} \notin \operatorname{span}(W_j)$, and we set

$$W_{j+1} = \begin{bmatrix} W_j & z_j^{(k+1)} \end{bmatrix} \quad \text{and} \quad t_{j+1} = t_j + 1. \tag{2.26}$$

Otherwise, it follows that $p_j^{(k+1)} \in \operatorname{span}(W_j)$, and we set

$$W_{j+1} = W_j \quad \text{and} \quad t_{j+1} = t_j. \tag{2.27}$$

At the end of the loop, we obtain $Z_{k+1} = W_{m+2}$ and $r_{k+1} = t_{m+2}$.

Exactly as in [4], the data obtained in the calculation of $Z_{k+1}$ can be used to compute

$\bar{g}_{k+1}$, $\bar{A}_{k+1}$ and $\bar{B}_{k+1}$ in a cheap way.

First, note that from (2.23), (2.24) and the fact that $s_k$, $g_k \in \text{span}(W_j)$, we have

$$(z_j^{(k+1)})^T p_j^{(k+1)} = \tau_j^{(k+1)}, \quad (z_j^{(k+1)})^T s_k = 0, \quad (z_j^{(k+1)})^T g_k = 0. \tag{2.28}$$

Now, by the calculation of $Z_{k+1}$, we have to consider two cases: $Z_{k+1} \neq Z_k$ or $Z_{k+1} = Z_k$. In the first case we have $Z_{k+1} = \begin{bmatrix} Z_k & \bar{Z}_{k+1} \end{bmatrix}$, then the Lemma 2.3 and the Remark 2.2 imply that $B_k \bar{Z}_{k+1} = \alpha \bar{Z}_{k+1}$ and the columns of $B_k Z_k$ belongs to $G_k$. Thus, denoting $t = r_{k+1} - r_k$ we get

$$\tilde{s}_k = Z_{k+1}^T s_k = \begin{bmatrix} Z_k^T s_k \\ \bar{Z}_{k+1}^T s_k \end{bmatrix} = \begin{bmatrix} \bar{s}_k \\ 0 \end{bmatrix} \tag{2.29}$$

$$\begin{aligned}
\tilde{B}_k = Z_{k+1}^T B_k Z_{k+1} &= \begin{bmatrix} Z_k^T \\ \bar{Z}_{k+1}^T \end{bmatrix} B_k \begin{bmatrix} Z_k & \bar{Z}_{k+1} \end{bmatrix} \\
&= \begin{bmatrix} Z_k^T \\ \bar{Z}_{k+1}^T \end{bmatrix} \begin{bmatrix} B_k Z_k & B_k \bar{Z}_{k+1} \end{bmatrix} = \begin{bmatrix} Z_k^T \\ \bar{Z}_{k+1}^T \end{bmatrix} \begin{bmatrix} B_k Z_k & \alpha \bar{Z}_{k+1} \end{bmatrix} \\
&= \begin{bmatrix} Z_k^T B_k Z_k & \alpha Z_k^T \bar{Z}_{k+1} \\ \bar{Z}_{k+1}^T B_k Z_k & \alpha \bar{Z}_{k+1}^T \bar{Z}_{k+1} \end{bmatrix} = \begin{bmatrix} \bar{B}_k & 0 \\ 0 & \alpha I_t \end{bmatrix}.
\end{aligned} \tag{2.30}$$

To compute $\bar{g}_{k+1}$, first note that

$$\bar{g}_{k+1} = Z_{k+1}^T g_{k+1} = \begin{bmatrix} Z_k^T g_{k+1} \\ \bar{Z}_{k+1}^T g_{k+1} \end{bmatrix}. \tag{2.31}$$

From (2.22) and (2.24), we have

$$\begin{aligned}
W_{m+1}^T p_{m+1}^{(k+1)} = u_{m+1}^{(k)} \Rightarrow W_{m+1}^T g_{k+1} &= u_{m+1}^{(k)} \\
\Rightarrow \begin{bmatrix} Z_k & \tilde{Z}_{k+1} \end{bmatrix}^T g_{k+1} &= u_{m+1}^{(k)} \\
\Rightarrow Z_k^T g_{k+1} &= \begin{bmatrix} \left(u_{m+1}^{(k)}\right)_1 & \cdots & \left(u_{m+1}^{(k)}\right)_{r_k} \end{bmatrix}^T,
\end{aligned} \tag{2.32}$$

where the columns of $\tilde{Z}_{k+1}$ are distinct vectors of the set $\left\{z_1^{(k+1)}, \ldots, z_m^{(k+1)}\right\}$. Moreover,

$$\begin{aligned}
\bar{Z}_{k+1}^T W_{m+1} &= \bar{Z}_{k+1}^T \left[\begin{array}{cc} Z_k & \tilde{Z}_{k+1} \end{array}\right] \\
&= \left[\begin{array}{cc} 0 & \bar{Z}_{k+1}^T \tilde{Z}_{k+1} \end{array}\right] \\
&= \begin{cases} \left[\begin{array}{c|c} 0 & I_{t-1} \\ \hline 0\ldots 0 & 0\ldots 0 \end{array}\right], & \text{if } \tau_{m+1}^{(k+1)} > 0, \\ \\ \left[\begin{array}{cc} 0 & I_t \end{array}\right], & \text{otherwise.} \end{cases}
\end{aligned} \tag{2.33}$$

Then, multiplying (2.23) for the left by $\bar{Z}_{k+1}^T$ (with $j = m+1$), we get

$$\begin{aligned}
\bar{Z}_{k+1}^T g_{k+1} &= \bar{Z}_{k+1}^T W_{m+1} u_{m+1}^{(k)} + \tau_{m+1}^{(k+1)} \bar{Z}_{k+1}^T z_{m+1}^{(k+1)} \\
&= \begin{cases} \left[\begin{array}{cccc} \left(u_{m+1}^{(k)}\right)_{r_k+1} & \cdots & \left(u_{m+1}^{(k)}\right)_{r_{k+1}-1} & \tau_{m+1}^{(k+1)} \end{array}\right]^T, & \text{if } \tau_{m+1}^{(k+1)} > 0, \\ \\ \left[\begin{array}{ccc} \left(u_{m+1}^{(k)}\right)_{r_k+1} & \cdots & \left(u_{m+1}^{(k)}\right)_{r_{k+1}} \end{array}\right]^T, & \text{otherwise.} \end{cases}
\end{aligned} \tag{2.34}$$

Thus, combining (2.31), (2.32) and (2.34), it follows that

$$\bar{g}_{k+1} = \begin{cases} \left[\begin{array}{cccc} \left(u_{m+1}^{(k)}\right)_1 & \cdots & \left(u_{m+1}^{(k)}\right)_{r_{k+1}-1} & \tau_{m+1}^{(k+1)} \end{array}\right]^T, & \text{if } \tau_{m+1}^{(k+1)} > 0, \\ \\ \left[\begin{array}{ccc} \left(u_{m+1}^{(k)}\right)_1 & \cdots & \left(u_{m+1}^{(k)}\right)_{r_{k+1}} \end{array}\right]^T, & \text{otherwise.} \end{cases} \tag{2.35}$$

Now, we will compute the matrix $\bar{A}_{k+1}$. By (2.22),

$$\begin{aligned}
\bar{A}_{k+1} = A_{k+1} Z_{k+1} &= \left[\begin{array}{cc} A_{k+1} Z_k & A_{k+1} \bar{Z}_{k+1} \end{array}\right] \\
&= \left[\begin{array}{cc} \left[\begin{array}{c} \left(p_1^{(k+1)}\right)^T Z_k \\ \vdots \\ \left(p_m^{(k+1)}\right)^T Z_k \end{array}\right] & \left[\begin{array}{c} \left(p_1^{(k+1)}\right)^T \bar{Z}_{k+1} \\ \vdots \\ \left(p_m^{(k+1)}\right)^T \bar{Z}_{k+1} \end{array}\right] \end{array}\right].
\end{aligned} \tag{2.36}$$

Thus, denoting

$$\bar{U}_{k+1} = \begin{bmatrix} \left(p_1^{(k+1)}\right)^T Z_k \\ \vdots \\ \left(p_m^{(k+1)}\right)^T Z_k \end{bmatrix} \tag{2.37}$$

and

$$\tilde{U}_{k+1} = \begin{bmatrix} \left(p_1^{(k+1)}\right)^T \bar{Z}_{k+1} \\ \vdots \\ \left(p_m^{(k+1)}\right)^T \bar{Z}_{k+1} \end{bmatrix} \tag{2.38}$$

it follows that

$$\bar{A}_{k+1} = \begin{bmatrix} \bar{U}_{k+1} & \tilde{U}_{k+1} \end{bmatrix}. \tag{2.39}$$

By (2.23), for each $j = 1, ..., m$,

$$\begin{aligned} & p_j^{(k+1)} = W_j u_j^{(k)} + \tau_j^{(k+1)} z_j^{(k+1)} \\ & \Longrightarrow \left(p_j^{(k+1)}\right)^T W_j = \left(u_j^{(k)}\right)^T W_j^T W_j + \tau_j^{(k+1)} \left(z_j^{(k+1)}\right)^T W_j \\ & \Longrightarrow \left(p_j^{(k+1)}\right)^T \begin{bmatrix} Z_k & \tilde{Z}_{k+1}^j \end{bmatrix} = \left(u_j^{(k)}\right)^T \\ & \Longrightarrow \left(p_j^{(k+1)}\right)^T Z_k = \begin{bmatrix} \left(u_j^{(k)}\right)_1 & \cdots & \left(u_j^{(k)}\right)_{r_k} \end{bmatrix}, \end{aligned} \tag{2.40}$$

where the columns of $\tilde{Z}_{k+1}^j$ are distinct vectors of the set $\left\{z_1^{(k+1)}, \ldots, z_{j-1}^{(k+1)}\right\}$. Further, taking the transposed of the equation (2.23) and multiplying for the left by $\bar{Z}_{k+1}$, we obtain

$$\left(p_j^{(k+1)}\right)^T \bar{Z}_{k+1} = \left(u_j^{(k)}\right)^T W_j^T \bar{Z}_{k+1} + \tau_j^{(k+1)} \left(z_j^{(k+1)}\right)^T \bar{Z}_{k+1}$$
$$= \left(u_j^{(k)}\right)^T \begin{bmatrix} Z_k & \tilde{Z}_{k+1}^j \end{bmatrix}^T \bar{Z}_{k+1} + \tau_j^{(k+1)} \left(z_j^{(k+1)}\right)^T \bar{Z}_{k+1}$$
$$= \begin{cases} \left(u_j^{(k)}\right)^T \left[\begin{array}{c|c} 0 & 0 \\ \hline I_{t_j - r_k} & 0 \end{array}\right] + \tau_j^{(k+1)} \begin{pmatrix} 0 & \cdots & 1 & 0 & \ldots & 0 \end{pmatrix}, & \text{if } \tau_j^{(k+1)} > 0 \\ \\ \left(u_j^{(k)}\right)^T \left[\begin{array}{c|c} 0 & 0 \\ \hline I_{t_j - r_k} & 0 \end{array}\right], & \text{otherwise} \end{cases}$$
$$= \begin{cases} \left[ \left(u_j^{(k)}\right)_{r_k+1} \cdots \left(u_j^{(k)}\right)_{t_j} \quad \tau_j^{(k+1)} \quad 0 \ \cdots \ 0 \right], & \text{if } \tau_j^{(k+1)} > 0 \\ \\ \left[ \left(u_j^{(k)}\right)_{r_k+1} \cdots \left(u_j^{(k)}\right)_{t_j} \quad 0 \ \cdots \ 0 \right], & \text{otherwise} \end{cases} \tag{2.41}$$

for each $j = 1, ..., m$, which completes the computation of $\bar{A}_{k+1}$.

Finally, if $y_k = (g_{k+1} - g_k) - \left(A_{k+1}^T \lambda_k - A_k^T \lambda_k\right)$, then

$$\tilde{y}_k = Z_{k+1}^T y_k = \begin{bmatrix} Z_k^T y_k \\ \\ \bar{Z}_{k+1}^T y_k \end{bmatrix}$$
$$= \begin{bmatrix} Z_k^T \left[g_{k+1} - g_k - A_{k+1}^T \lambda_k + A_k^T \lambda_k\right] \\ \\ \bar{Z}_{k+1}^T \left[g_{k+1} - g_k - A_{k+1}^T \lambda_k + A_k^T \lambda_k\right] \end{bmatrix}$$
$$= \begin{bmatrix} Z_k^T g_{k+1} - \bar{g}_k - \bar{U}_{k+1}^T \lambda_k + \bar{A}_k^T \lambda_k \\ \\ \bar{Z}_{k+1}^T g_{k+1} - \tilde{U}_{k+1}^T \lambda_k \end{bmatrix}. \tag{2.42}$$

Now, considering the case in which $Z_{k+1} = Z_k$, it follows that

$$\tilde{s}_k = Z_k^T s_k = \bar{s}_k, \tag{2.43}$$
$$\tilde{B}_k = Z_k^T B_k Z_k = \bar{B}_k, \tag{2.44}$$
$$\bar{g}_{k+1} = Z_k^T g_{k+1} = \left[ \left(u_{m+1}^{(k)}\right)_1 \ \cdots \ \left(u_{m+1}^{(k)}\right)_{r_k} \right]^T, \tag{2.45}$$
$$\bar{A}_{k+1} = A_{k+1} Z_k = \bar{U}_{k+1}, \tag{2.46}$$
$$\tilde{y}_k = Z_k^T y_k = \bar{g}_{k+1} - \bar{g}_k - \bar{U}_{k+1}^T \lambda_k + \bar{A}_k^T \lambda_k. \tag{2.47}$$

By Lemma 2.5, the reduced matrix

$$\bar{B}_{k+1} = Z_{k+1}^T B_{k+1} Z_{k+1}$$

in the subspace span$(Z_{k+1})$ can be obtained by any formula among the PSB and the Broyden family, by using $\tilde{s}_k$, $\tilde{B}_k$ and $\tilde{y}_k$ computed by (2.29), (2.30) and (2.42), or by (2.43), (2.44) and (2.47). Then, by Theorem 2.4 we can solve the subproblem (2.4)-(2.5) with the reduced matrices $\bar{B}_{k+1}$, $\bar{A}_{k+1}$ and reduced gradient $\bar{g}_{k+1}$ to obtain $\bar{s}_{k+1}$ and the trial step $s_{k+1} = Z_{k+1}\bar{s}_{k+1}$.

Now, we can summarize the main steps of our subspace version of the ALTR method.

**Algorithm 2.1.** *Subspace version of the Augmented Lagrangian Trust Region algorithm (ALTR)*

***Step 0*** *Given $x_1 \in \mathbb{R}^n$, $\sigma_1 > 1$, $\delta_1 > 0$, $\lambda_1 \in \mathbb{R}^m$, $\lambda_{min} < \lambda_{max}$, $\theta > 1$, $\beta \in (0,1)$, $\Delta_1 > 0$, $0 < \eta < \eta_1 < \frac{1}{2}$ and $R_1 = \max\{\|c(x_1)\|_2, 1\}$. Choose one quasi-Newton formula among PSB and Broyden family. Compute $\nabla c_1(x_1), \ldots, \nabla c_m(x_1)$ and $g_1 = \nabla f(x_1)$. Apply the Gram-Schmidt procedure with reorthogonalization to the set*

$$\{\nabla c_1(x_1), \ldots, \nabla c_m(x_1), g_1\}$$

*in order to obtain a column orthogonal matrix, $Z_1 \in \mathbb{R}^{n\times r_1}$, such that*

$$\operatorname{span}(Z_1) = \operatorname{span}\{\nabla c_1(x_1), \ldots, \nabla c_m(x_1), g_1\}. \tag{2.48}$$

*Set $\bar{B}_1 = \alpha I_{r_1}$, $\bar{g}_1 = Z_1^T g_1$, $\bar{A}_1 = A_1 Z_1$ and $k := 1$.*

***Step 1*** *If $\|c_k\|_2 = 0$ and $P_{\mathrm{Null}(\bar{A}_k)}(\bar{g}_k) = 0$, then stop.*

***Step 2*** *Set $\sigma_k^1 = \sigma_k$ and $l := 1$*

*While $\nabla_x L(x_k, \lambda_k; \sigma_k^l) = 0$ and $\|c_k\|_2 > 0$,*

$$\sigma_k^{l+1} := \theta\sigma_k^l \quad \text{and} \quad l := l+1 \tag{2.49}$$

*End (While)*

*Set $\sigma_k = \sigma_k^l$. Compute $\bar{s}_k$ by solving (2.4)-(2.5).*

***Step 3*** *Compute $s_k = Z_k\bar{s}_k$ and $\rho_k$ by (1.10). If $\rho_k \geq \eta$, go to Step 4. Otherwise, set $\Delta_{k+1} = \frac{\|s_k\|_2}{4}$, $x_{k+1} = x_k$, $k := k+1$ and go back to Step 2.*

***Step 4*** *If (1.11) is satisfied, then set*

$$\sigma_{k+1} = 2\sigma_k \quad \text{and} \quad \delta_{k+1} = \frac{\delta_k}{4}. \tag{2.50}$$

*Otherwise, set $\sigma_{k+1} = \sigma_k$ and $\delta_{k+1} = \delta_k$. If $\|c_{k+1}\|_2 \leq R_k$, compute $\lambda_{k+1}$ by (1.13)-(1.14) and set $R_{k+1} = \beta R_k$.*

*Otherwise, set $\lambda_{k+1} = \lambda_k$ and $R_{k+1} = R_k$.*

***Step 5*** *Set* $x_{k+1} = x_k + s_k$ *and*

$$\Delta_{k+1} = \begin{cases} \max\left\{\Delta_k, 1.5\left\|s_k\right\|_2\right\}, & \textit{if } \rho_k \in [1-\eta_1, +\infty) \ , \\ \Delta_k, & \textit{if } \rho_k \in [\eta_1, 1-\eta_1), \\ \max\left\{0.5\Delta_k, 0.75\left\|s_k\right\|_2\right\}, & \textit{if } \rho_k \in [\eta, \eta_1); \end{cases}$$

***Step 6*** *Calculate* $f_{k+1}$, $g_{k+1}$, $c_{k+1}$ *and* $A_{k+1}$. *If* $r_k = n$, *set* $\bar{A}_{k+1} = A_{k+1}$, $\bar{g}_{k+1} = g_{k+1}$, $\tilde{s}_k = s_k$, $\tilde{B}_k = \bar{B}_k$, $\tilde{y}_k = (g_{k+1} - g_k) - (A_{k+1} - A_k)^T \lambda_k$, $Z_{k+1} = I_n$, $r_{k+1} = n$ *and go to Step 8.*

***Step 7*** *Set* $W_1 = Z_k$, $t_1 = r_k$, *and consider the notation (2.22).*
*For* $j = 1 : m+1$
*(a) Obtain (2.23) by the reorthogonalization procedure;*
*(b) If* $\tau_j^{(k+1)} > 0$, *set* $W_{j+1} = \begin{bmatrix} W_j & z_j^{(k+1)} \end{bmatrix}$ *and* $t_{j+1} = t_j + 1$.
*Otherwise, set* $W_{j+1} = W_j$ *and* $t_{j+1} = t_j$.
*End(For)*
*Set* $Z_{k+1} = W_{m+2}$ *and* $r_{k+1} = t_{m+2}$;
*If* $Z_{k+1} \neq Z_k$ *compute* $\tilde{s}_k$, $\tilde{B}_k$, $\bar{g}_{k+1}$, $\bar{A}_{k+1}$, $\tilde{y}_k$ *according to (2.29), (2.30), (2.35), (2.39) and (2.42), respectively. Otherwise, compute* $\tilde{s}_k$, $\tilde{B}_k$, $\bar{g}_{k+1}$, $\bar{A}_{k+1}$, $\tilde{y}_k$ *by (2.43)-(2.47) respectively.*

***Step 8*** *Obtain* $\bar{B}_{k+1} = Update(\tilde{B}_k, \tilde{s}_k, \tilde{y}_k)$ *by the chosen matrix updating formula.*
*Set* $k := k+1$ *and go back to Step 1.*

**Remark 2.4.** *By Step 6, when the dimension* $r_k$ *of the subspace* $\operatorname{span}(Z_k)$ *reaches* $n$, *the subspace ALTR reduces to the standard ALTR. The reason for this step is to avoid the computational effort required by Step 7 when it is not necessary anymore.*

# Chapter 3

# Global Convergence

If $\text{span}(Z_k) = G_k$ and we solve (2.4)-(2.5) exactly at all iterations, it follows from Theorem 2.4 that Algorithm 2.1 is equivalent to Algorithm 1.1. Even if (2.4)-(2.5) is solved inexactly, this equivalence also happens when $r_k$ reaches $n$, because then we have $\text{span}(Z_k) = R^n$ for all subsequent iterations. In both cases, the global convergence of Algorithm 2.1 follows from the fact that Algorithm 1.1 (the original ALTR) is globally convergent (see Sections 3.1 and 3.2 in [13]). In this chapter, we will study global convergence properties of Algorithm 2.1 in a more general setting that allows $\text{span}(Z_k) \neq G_k$ and the inexact solution of (2.4)-(2.5), simultaneously[1]. Specifically, we consider the following assumptions:

**A1** The functions $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ and $\nabla c_i : \mathbb{R}^n \to \mathbb{R}^n$ $(i = 1, \ldots, m)$ are Lipschitz;

**A2** The sequences $\{x_k\}$ and $\{\|\bar{B}_k\|_2\}$ are bounded;

**A3** For each $k$, $Z_k^T Z_k = I_{r_k}$, $\{\nabla c_1(x_k), \ldots, \nabla c_m(x_k), g_k\} \subset \text{span}(Z_k)$ and $B_k z \in \text{span}(Z_k)$ for all $z \in \text{span}(Z_k)$;

**A4** For all $k$, the approximate solution $\bar{s}_k$ satisfies

$$\bar{q}_k(0) - \bar{q}_k(\bar{s}_k) \geq \beta \left[\bar{q}_k(0) - \bar{q}_k(\bar{s}_k^c)\right]$$

for some $\beta \in (0, 1)$, where $\bar{s}_k^c$ is the Cauchy step to (2.4)-(2.5).

It is useful to consider the following remark, which will be extensively called in the proofs.

**Remark 3.1.** *From $Z_k^T Z_k = I_{r_k}$, it follows that*

$$v \in \text{span}(Z_k) \;\Rightarrow\; v = Z_k Z_k^T v. \tag{3.1}$$

**Lemma 3.1.** *Suppose that A1-A3 hold. Then, the sequence $\{\|\bar{A}_k\|_2\}$ is bounded.*

[1]Our analysis allows choices to $Z_{k+1}$ different from that described at Step 7 in Algorithm 2.1.

*Proof:* By A1 and A2 there exists $\kappa_1 > 0$ such that

$$\|A_k\|_2 \leq \kappa_1, \text{ for all } k. \tag{3.2}$$

On the other hand, given $x \in \mathbb{R}^m$ and noting that

$$A_k^T = [\nabla c_1(x_k) \ldots \nabla c_m(x_k)], \tag{3.3}$$

the condition A3 implies that $A_k^T x \in \text{span}(Z_k)$ and, by Remark 3.1 it follows that

$$\begin{aligned} \|\bar{A}_k^T x\|_2^2 &= \|Z_k^T A_k^T x\|_2^2 \\ &= (Z_k^T A_k^T x)^T (Z_k^T A_k^T x) \\ &= (A_k^T x)^T Z_k Z_k^T A_k^T x \\ &= (A_k^T x)^T A_k^T x \\ &= \|A_k^T x\|_2^2 \end{aligned}$$

$$\Longrightarrow \|\bar{A}_k^T x\|_2 = \|A_k^T x\|_2 \tag{3.4}$$

Hence,

$$\|\bar{A}_k^T\|_2 = \max_{\|x\|_2=1} \|\bar{A}_k^T x\|_2 = \max_{\|x\|_2=1} \|A_k^T x\|_2 = \|A_k^T\|_2 \leq \kappa_1, \text{ for all } k. \tag{3.5}$$

■

**Lemma 3.2.** *Suppose that A1-A3 hold and $A_k^T$ has full column rank. Then we have*

$$\left\|P_{\text{Null}(\bar{A}_k)}\left(\bar{g}_k - \bar{A}_k^T \lambda_k\right)\right\|_2 = \left\|P_{\text{Null}(A_k)}\left(g_k - A_k^T \lambda_k\right)\right\|_2, \tag{3.6}$$

*where $P_{\text{Null}(A_k)}$ is the orthogonal projection matrix onto null space of $A_k$ and $P_{\text{Null}(\bar{A}_k)}$ is the orthogonal projection matrix onto null space of $\bar{A}_k$.*

*Proof:* It is known that the orthogonal projection onto the subspace $\text{Range}(A_k^T)$ is defined as $A_k^T(A_k^T)^\dagger$. Since $A_k^T$ has full column rank, then $(A_k^T)^\dagger = ((A_k^T)^T A_k^T)^{-1}(A_k^T)^T = (A_k A_k^T)^{-1} A_k$ (see, e.g., page 257 in [3]). It is also known that $\text{Null}(A_k) = \text{Range}(A_k^T)^\perp$, hence the orthogonal projection matrix onto the subspace $\text{Null}(A_k)$ is defined by (see page 75 in [3]):

$$P_{\text{Null}(A_k)} = I_n - A_k^T(A_k^T)^\dagger = I_n - A_k^T(A_k A_k^T)^{-1} A_k. \tag{3.7}$$

Thus, it follows that

$$
\begin{aligned}
P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k) &= (I_n - A_k^T(A_kA_k^T)^{-1}A_k)(g_k - A_k^T\lambda_k) \\
&= g_k - A_k^T\lambda_k - A_k^T(A_kA_k^T)^{-1}A_kg_k + A_k^T(A_kA_k^T)^{-1}A_kA_k^T\lambda_k \\
&= g_k - A_k^T(A_kA_k^T)^{-1}A_kg_k. \qquad (3.8)
\end{aligned}
$$

On the other hand, the matrix $\bar{A}_k^T = Z_k^TA_k^T$ has also full column rank, because otherwise, there would be scalars $\alpha_i$, $i = 1, ..., m$, with at least one of them being nonzero, such that

$$\sum_{i=1}^{m} \alpha_i Z_k^T \nabla c_i(x_k) = 0. \qquad (3.9)$$

Multiplying both sides of the equation (3.9) by $Z_k$, from A3 and Remark 3.1 it follows that

$$\sum_{i=1}^{m} \alpha_i \nabla c_i(x_k) = 0, \qquad (3.10)$$

with at least one $\alpha_i$ nonzero. In this way, the columns of the matrix $A_k^T$ would form a linearly dependent set, which contradicts the fact that $A_k^T$ has full column rank, therefore the last statement is true. Thus, similarly to the previous discussion, the orthogonal projection matrix onto the subspace $\mathrm{Null}(\bar{A}_k)$ is defined by

$$P_{\mathrm{Null}(\bar{A}_k)} = I_{r_k} - \bar{A}_k^T(\bar{A}_k^T)^{\dagger} = I_{r_k} - \bar{A}_k^T(\bar{A}_k\bar{A}_k^T)^{-1}\bar{A}_k. \qquad (3.11)$$

By this, A3, Remark 3.1 and (3.8) we have

$$
\begin{aligned}
P_{\mathrm{Null}(\bar{A}_k)}(\bar{g}_k - \bar{A}_k^T\lambda_k) &= \bar{g}_k - \bar{A}_k^T\lambda_k - \bar{A}_k^T(\bar{A}_k\bar{A}_k^T)^{-1}\bar{A}_k\bar{g}_k + \bar{A}_k^T(\bar{A}_k\bar{A}_k^T)^{-1}\bar{A}_k\bar{A}_k^T\lambda_k \\
&= Z_k^Tg_k - Z_k^TA_k^T(A_kZ_kZ_k^TA_k^T)^{-1}A_kZ_kZ_k^Tg_k \\
&= Z_k^Tg_k - Z_k^TA_k^T(A_kA_k^T)^{-1}A_kg_k \\
&= Z_k^T(g_k - A_k^T(A_kA_k^T)^{-1}A_kg_k) \\
&= Z_k^T(P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k)). \qquad (3.12)
\end{aligned}
$$

Now, note that equation (3.8) implies $P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k) \in \mathrm{span}(Z_k)$. Hence, by (3.12) and Remark 3.1 we get

$$
\begin{aligned}
\left\|P_{\mathrm{Null}(\bar{A}_k)}(\bar{g}_k - \bar{A}_k^T\lambda_k)\right\|_2^2 &= \left\|Z_k^T(P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k))\right\|_2^2 \\
&= (P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k))^TZ_kZ_k^T(P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k)) \\
&= \left\|(P_{\mathrm{Null}(A_k)}(g_k - A_k^T\lambda_k))\right\|_2^2.
\end{aligned}
$$

This completes the proof. ∎

**Lemma 3.3.** *There is a constant $\xi > 0$ such that the inequality*

$$Pred_k \geq \bar{\beta}\frac{1}{2}\|\hat{g}_k\|_2 \min\left\{\frac{\|\hat{g}_k\|_2}{\sigma_k \xi}, \Delta_k\right\} \tag{3.13}$$

*holds for all $k$, where $\hat{g}_k = g_k - A_k^T\lambda_k + \sigma_k A_k^T c_k$.*

*Proof:* We have

$$\begin{aligned}
\overline{Pred}_k &= \bar{q}_k(0) - \bar{q}_k(\bar{s}_k) \\
&= \frac{\sigma_k}{2}\|c_k\|_2^2 - (\bar{g}_k^T\bar{s}_k - \lambda_k^T\bar{A}_k\bar{s}_k + \frac{1}{2}\bar{s}_k^T\bar{B}_k\bar{s}_k + \frac{\sigma_k}{2}\left\|c_k + \bar{A}_k\bar{s}_k\right\|_2^2) \\
&= \frac{\sigma_k}{2}\|c_k\|_2^2 - (g_k^T Z_k Z_k^T s_k - \lambda_k^T A_k Z_k Z_k^T s_k + \frac{1}{2}s_k^T Z_k Z_k^T B_k Z_k Z_k^T s_k \\
&\quad + \frac{\sigma_k}{2}\left\|c_k + A_k Z_k Z_k^T s_k\right\|_2^2)
\end{aligned} \tag{3.14}$$

Since $s_k \in \operatorname{span}(Z_k)$, from Remark 3.1 and A3 it follows that $s_k = Z_k Z_k^T s_k$, $B_k s_k \in \operatorname{span}(Z_k)$ and $B_k s_k = Z_k Z_k^T B_k s_k$. Then, by (3.14) we get

$$\begin{aligned}
\overline{Pred}_k &= \frac{\sigma_k}{2}\|c_k\|_2^2 - (g_k^T s_k - \lambda_k^T A_k s_k + \frac{1}{2}s_k^T B_k s_k + \frac{\sigma_k}{2}\|c_k + A_k s_k\|_2^2) \\
&= q_k(0) - q_k(s_k) \\
&= Pred_k.
\end{aligned} \tag{3.15}$$

Now, note that the subproblem (2.4)-(2.5) can be rewritten as follows

$$\min_{\bar{s}\in\mathbb{R}^r} \ \bar{q}_k(\bar{s}) \equiv \breve{g}_k^T\bar{s} + \frac{1}{2}\bar{s}^T(\bar{B}_k + \sigma_k\bar{A}_k^T\bar{A}_k)\bar{s} + \frac{\sigma_k}{2}\|c_k\|_2^2 \tag{3.16}$$

$$\text{s. t.} \ \ \|\bar{s}\|_2 \leq \Delta_k, \tag{3.17}$$

where $\breve{g}_k = \bar{g}_k - \bar{A}_k^T\lambda_k + \sigma_k\bar{A}_k^T c_k$. By A4 and Lemma 5.36 in [9], we have

$$\bar{q}_k(0) - \bar{q}_k(\bar{s}_k) = \overline{Pred}_k \geq \bar{\beta}\frac{1}{2}\|\breve{g}_k\|_2 \min\left\{\frac{\|\breve{g}_k\|_2}{\left\|\bar{B}_k + \sigma_k\bar{A}_k^T\bar{A}_k\right\|_2}, \Delta_k\right\}, \tag{3.18}$$

for some $\bar{\beta} \in (0,1)$. Note that

$$\begin{aligned}
\breve{g}_k &= \bar{g}_k - \bar{A}_k^T\lambda_k + \sigma_k\bar{A}_k^T c_k \\
&= Z_k^T g_k - Z_k^T A_k^T\lambda_k + \sigma_k Z_k^T A_k^T c_k \\
&= Z_k^T(g_k - A_k^T\lambda_k + \sigma_k A_k^T c_k) \\
&= Z_k^T\hat{g}_k.
\end{aligned} \tag{3.19}$$

By A3, $\hat{g}_k \in \operatorname{span}(Z_k)$. Thus, Remark 3.1 implies $\hat{g}_k = Z_k Z_k^T \hat{g}_k$ and, consequently,

$$\begin{aligned}
\|\breve{g}_k\|_2^2 &= \left\|Z_k^T \hat{g}_k\right\|_2^2 \\
&= \hat{g}_k^T Z_k Z_k^T \hat{g}_k \\
&= \hat{g}_k^T \hat{g}_k \\
&= \|\hat{g}_k\|_2^2 \\
\Longrightarrow \|\breve{g}_k\|_2 &= \|\hat{g}_k\|_2 . \qquad (3.20)
\end{aligned}$$

From A2 and Lemma 3.1, there are positive constants $\xi_1$ and $\kappa_2$ such that

$$\left\|\bar{B}_k + \sigma_k \bar{A}_k^T \bar{A}_k\right\|_2 \leq \xi_1 + \sigma_k \kappa_2 \leq \sigma_k \xi_1 + \sigma_k \kappa_2 = \sigma_k(\xi_1 + \kappa_2) = \sigma_k \xi, \qquad (3.21)$$

where $\xi = \xi_1 + \kappa_2$. Therefore, from (3.15), (3.18), (3.20) and (3.21) we get (3.13). ∎

The following remark will be used in the next proofs.

**Remark 3.2.** *By the calculation of the vector $\lambda_k$ and by A1 and A2 it follows that*

$$\begin{aligned}
|f(x_k)| &\leq f_{\max} \quad \forall k \qquad (3.22) \\
\|c(x_k)\|_2 &\leq \|c_{\max}\|_2 \quad \forall k \qquad (3.23) \\
\|\lambda_k\|_2 &\leq \|\lambda_{\max}\|_2 \quad \forall k. \qquad (3.24)
\end{aligned}$$

By the construction of the subspace version of ALTR algorithm, as well as of ALTR itself, we must consider two cases: when the penalty parameters $\sigma_k$ increase indefinitely and when the sequence $\{\sigma_k\}$ is bounded. The analysis is done by adapting the corresponding proofs in [13].

## 3.1 Case I: Sequence of penalty parameters is unbounded

The first result shows that when the penalty parameters increase to infinity, we find that the sequence of constraint violations $\{\|c_k\|_2\}$ is convergent. Its proof is due to Wang and Yuan [13]. We give it here for completeness.

**Lemma 3.4.** *Suppose that A1-A2 hold. If $\sigma_k \to \infty$, then $\lim_{k\to\infty} \|c_k\|_2$ exists.*

*Proof:* We have that the penalty parameters $\sigma_k$, $k = 1, 2, 3, \ldots$, are positive and increase monotonically. Therefore, the numbers $\sigma_i^{-1} - \sigma_{i+1}^{-1}$, $i \geq 1$, are all nonnegatives, and, for all integers $p$ and $q$ that satisfy $0 < p < q$, their sum has the property

$$\sum_{i=p}^{q-1}(\sigma_i^{-1} - \sigma_{i+1}^{-1}) = \sigma_p^{-1} - \sigma_q^{-1}. \tag{3.25}$$

From this and Remark 3.2 we have

$$\begin{aligned}
\sum_{i=p}^{q} \sigma_i^{-1}\left[f(x_i) - f(x_{i+1})\right] &= \sigma_p^{-1} f(x_p) + \sum_{i=p}^{q-1}(-\sigma_i^{-1} + \sigma_{i+1}^{-1}) f(x_{i+1}) - \sigma_q^{-1} f(x_{q+1}) \\
&\leq \left|\sigma_p^{-1} f(x_p)\right| + \sum_{i=p}^{q-1}\left|-\sigma_i^{-1} + \sigma_{i+1}^{-1}\right| \left|f(x_{i+1})\right| + \left|\sigma_q^{-1} f(x_{q+1})\right| \\
&\leq \sigma_p^{-1} f_{\max} + (\sigma_p^{-1} - \sigma_q^{-1}) f_{\max} + \sigma_q^{-1} f_{\max} \\
&= 2\sigma_p^{-1} f_{\max}.
\end{aligned} \tag{3.26}$$

We also have the bound for the sum

$$\begin{aligned}
\sum_{i=p}^{q} \sigma_i^{-1}\left[\lambda_i^T c(x_{i+1}) - \lambda_i^T c(x_i)\right] = &- \sigma_p^{-1}\lambda_p^T c(x_p) + \sigma_q^{-1}\lambda_q^T c(x_{q+1}) \\
&+ \sum_{i=p}^{q-1}\left\{\sigma_i^{-1}\lambda_i^T c(x_{i+1}) - \sigma_{i+1}^{-1}\lambda_{i+1}^T c(x_{i+1})\right\}.
\end{aligned} \tag{3.27}$$

Let the set $\mathcal{I}(p,q)$ of the integers numbers in the interval $[p, q-1]$ such that $\lambda_{i+1} \neq \lambda_i$, but this set can be empty. From the Cauchy-Schwarz inequality, Remark 3.2, (3.25) and update rule to $\lambda_i$ we get

$$\begin{aligned}
\sum_{i\in\mathcal{I}(p,q)} \left\{\sigma_i^{-1}\lambda_i^T c(x_{i+1}) - \sigma_{i+1}^{-1}\lambda_{i+1}^T c(x_{i+1})\right\} &\leq \sum_{i\in\mathcal{I}(p,q)} \left\{\sigma_i^{-1}\|\lambda_i\|_2 + \sigma_{i+1}^{-1}\|\lambda_{i+1}\|_2\right\}\|c(x_{i+1})\|_2 \\
&\leq \|\lambda_{\max}\|_2 \sum_{i\in\mathcal{I}(p,q)} (\sigma_i^{-1} + \sigma_{i+1}^{-1}) R_i \\
&= \|\lambda_{\max}\|_2 \sum_{i\in\mathcal{I}(p,q)} (\sigma_i^{-1} + \sigma_{i+1}^{-1}) \beta^i R_0 \\
&\leq \|\lambda_{\max}\|_2 \sum_{i=p}^{q-1} (\sigma_i^{-1} + \sigma_{i+1}^{-1}) \beta^i R_0 \\
&\leq \|\lambda_{\max}\|_2 \sum_{i=p}^{q-1} (\sigma_i^{-1} + \sigma_{i+1}^{-1}) \sum_{i=p}^{q-1} \beta^i R_0 \\
&\leq 2\sigma_p^{-1}\|\lambda_{\max}\|_2 \sum_{i=1}^{\infty} \beta^i R_0 \\
&= 2\sigma_p^{-1}\|\lambda_{\max}\|_2 \frac{R_0}{(1-\beta)},
\end{aligned} \tag{3.28}$$

where we used the fact that $\beta \in (0,1)$ and $\sigma_p \le \sigma_i$ for all $i > p \Rightarrow \sigma_p^{-1} \ge \sigma_i^{-1}$ for all $i > p$. For all the other terms in the second line of (3.27), the vectors $\lambda_i$ and $\lambda_{i+1}$ are the same, then it follows

$$\begin{aligned}\sum_{i \notin \mathcal{I}(p,q)} \left\{\sigma_i^{-1}\lambda_i^T c(x_{i+1}) - \sigma_{i+1}^{-1}\lambda_{i+1}^T c(x_{i+1})\right\} &= \sum_{i \notin \mathcal{I}(p,q)} \left(\sigma_i^{-1} - \sigma_{i+1}^{-1}\right)\lambda_i^T c(x_{i+1}) \\ &\le \|\lambda_{\max}\|_2 \|c_{\max}\|_2 \sum_{i=p}^{q-1} (\sigma_i^{-1} - \sigma_{i+1}^{-1}) \\ &= (\sigma_p^{-1} - \sigma_q^{-1}) \|\lambda_{\max}\|_2 \|c_{\max}\|_2 \end{aligned} \tag{3.29}$$

Now from (3.27)-(3.29) we obtain

$$\begin{aligned}\sum_{i=p}^{q} \sigma_i^{-1}\left[\lambda_i^T c(x_{i+1}) - \lambda_i^T c(x_i)\right] &\le \sigma_p^{-1}\|\lambda_{\max}\|_2 \|c_{\max}\|_2 + \sigma_q^{-1}\|\lambda_{\max}\|_2 \|c_{\max}\|_2 \\ &\quad + 2\sigma_p^{-1}\|\lambda_{\max}\|_2 \frac{R_0}{(1-\beta)} + (\sigma_p^{-1} - \sigma_q^{-1})\|\lambda_{\max}\|_2 \|c_{\max}\|_2 \\ &= 2\sigma_p^{-1}\|\lambda_{\max}\|_2 \left\{\|c_{\max}\|_2 + \frac{R_0}{(1-\beta)}\right\}. \end{aligned} \tag{3.30}$$

By the construction of Algorithm 2.1 we can see that $L(x_{k+1}, \lambda_i; \sigma_i) \le L(x_k, \lambda_i; \sigma_i)$ for all $k$, and $x_{k+1}$ is different from $x_k$ only if $Ared_k$ is positive, that is, for $k$ such that the trial step $s_k$ is accepted. Then, (3.26) and (3.30) imply that

$$\sum_{i=p}^{q} \sigma_i^{-1}\left\{L(x_i, \lambda_i, \sigma_i) - L(x_{i+1}, \lambda_i, \sigma_i)\right\} \le \frac{1}{2}\|c(x_p)\|_2^2 - \frac{1}{2}\|c(x_{q+1})\|_2^2 + M_0\sigma_p^{-1}, \tag{3.31}$$

where $M_0$ is a constant

$$M_0 = 2f_{\max} + 2\|\lambda_{\max}\|_2 \left\{\|c_{\max}\|_2 + \frac{R_0}{(1-\beta)}\right\}.$$

Note that the sum on the left-hand side of inequality (3.31) is bounded above by $M_0\sigma_p^{-1} + \frac{1}{2}\|c(x_p)\|_2^2$ independently of who is $q$. Hence, by letting $q \to \infty$ for any $p$ fixed, it follows that the sum of the products $\sigma_k^{-1}\left\{L(x_k, \lambda_k, \sigma_k) - L(x_{k+1}, \lambda_k, \sigma_k)\right\}, k = 1, 2, 3, ...$, is absolutely convergent. Further, the nonnegativity of the left-hand side of inequality (3.31) gives the condition

$$\|c(x_{q+1})\|_2 \le \|c(x_p)\|_2^2 + 2M_0\sigma_p^{-1}, \quad 0 < p < q. \tag{3.32}$$

By letting $q \to \infty$ again with $p$ fixed, we obtain that the sequence $\left\{\|c(x_k)\|_2^2\right\}$ is bounded. Denote the lim inf by $\|c_\infty\|_2^2$, and, for any $\epsilon > 0$ arbitrary, let $p$ satisfying $\|c(x_p)\|_2^2 < \|c_\infty\|_2^2 + \epsilon$. Since $\sigma_k \to \infty$, the choice of $p$ can also satisfy $2M_0\sigma_p^{-1} < \epsilon$. From condition

(3.32) it follows that

$$\|c(x_{q+1})\|_2 \leq \|c_\infty\|_2^2 + 2\epsilon, \quad q > p.$$

Therefore, the lim sup of sequence $\left\{\|c(x_k)\|_2^2\right\}$ when $k \to \infty$ is, at most, $\|c_\infty\|_2^2 + 2\epsilon$. Now, since the number $\epsilon$ can be arbitrarily small, we conclude that lim inf and lim sup of sequence $\left\{\|c(x_k)\|_2^2\right\}$ are the same. The proof is complete. ■

Lemma 3.4 ensures the convergence of $\{\|c(x_k)\|_2\}$ when $\sigma_k \to \infty$. Consequently, two cases arise: either all the accumulation points of $\{x_k\}$ are infeasible, or all the accumulation points are feasible. We are interested in methods that can find feasible accumulation points, but that is impossible if the original problem is "naturally" infeasible, for example, $c(x) \neq 0$ for any $x \in \mathbb{R}^n$. Therefore, it is necessary to analyze how the algorithm behaves in terms of infeasibility.

**Theorem 3.5.** *Suppose A1-A2 and A4 hold. If* $\lim_{k\to\infty} \sigma_k = \infty$ *and* $\lim_{k\to\infty} \|c_k\|_2 > 0$, *then any accumulation point of* $\{x_k\}$ *is a KKT point of the problem*

$$\min_{x\in\mathbb{R}^n} \|c(x)\|_2^2 . \tag{3.33}$$

*Proof:* A KKT point of (3.33) is characterized by $A(x)^T c(x) = 0$. Then to prove the theorem, we have to establish that $\lim_{k\to\infty} \left\|A_k^T c_k\right\|_2 = 0$. First, it will be prove that

$$\liminf_{k\to\infty} \left\|A_k^T c_k\right\|_2 = 0. \tag{3.34}$$

By contradiction, assume that (3.34) is not true. Then, there exists a constant $\mu > 0$ such that

$$\liminf_{k\to\infty} \left\|A_k^T c_k\right\|_2 \geq 2\mu. \tag{3.35}$$

As previously denoted, $\hat{g}_k = \nabla_x L(x_k, \lambda_k, \sigma_k) = g_k - A_k^T \lambda_k + \sigma_k A_k^T c_k$. Due to $\sigma_k \to \infty$, $\left\|g_k - A_k^T \lambda_k\right\|_2$ be bounded (by A1-A2 and the computation of $\lambda_k$) and (3.35), there exists $k_0$ such that the following property holds

$$\|\hat{g}_k\|_2 \geq \sigma_k \left\|A_k^T c_k\right\|_2 - \left\|g_k - A_k^T \lambda_k\right\|_2 \geq \mu\sigma_k, \quad \text{for all } \ k \geq k_0. \tag{3.36}$$

From (3.36) and Step 2 of Algorithm 2.1, it follows that the update (2.49) occurs only in a finite number of iterations. Without loss of generality, we assume (2.49) never happens. Then, the fact that $\sigma_k \to \infty$ is due to (2.50) happening infinitely many times. By Lemma 3.3 and (3.36), we have that the predict reduction obtained at $s_k$ satisfies the following inequality

$$Pred_k \geq \bar{\beta}\frac{1}{2}\|\hat{g}_k\|_2 \min\left\{\frac{\|\hat{g}_k\|_2}{\sigma_k\xi}, \Delta_k\right\} \geq \frac{\bar{\beta}\mu\sigma_k}{2}\min\left\{\frac{\mu}{\xi}, \Delta_k\right\}, \text{ for all } k \geq k_0. \tag{3.37}$$

However, $\sigma_k \to \infty$ provides the condition $\delta_k\sigma_k \to 0$. Further,

$$Pred_k < \delta_k\sigma_k \min\left\{\Delta_k \|c_k\|_2, \|c_k\|_2^2\right\}$$

holds for infinitely many $k$. This contradicts (3.37). Hence, (3.34) is true.

Now, by contradiction again, assume that $\lim_{k\to\infty}\left\|A_k^T c_k\right\|_2 = 0$ does not hold. Then there exists a subsequence such that

$$\left\|A_{t_i}^T c_{t_i}\right\|_2 \geq 2\epsilon \tag{3.38}$$

for some $\epsilon > 0$ and for all $i$ sufficiently large. Since $\sigma_k \to \infty$, then it follows that $\sigma_{t_i} \to \infty$ for $i$ sufficiently large. By (3.34), it is ensured that for each $i$, there is an index $l(t_i) > t_i$, where $l(t_i)$ is the first index bigger than $t_i$ such that $\left\|A_{l(t_i)}^T c_{l(t_i)}\right\|_2 < \epsilon$. Denote $l_i := l(t_i)$. Thus, there exists another subsequence $\{l_i\}$ such that

$$\left\|A_k^T c_k\right\|_2 \geq \epsilon \quad \text{for } t_i \leq k < l_i \quad \text{and} \quad \left\|A_{l_i}^T c_{l_i}\right\|_2 < \epsilon. \tag{3.39}$$

Let $\mathcal{L} := \bigcup_i \{k;\ t_i \leq k < l_i\}$. From $\{\sigma_k\}_{\mathcal{L}} \to \infty$, $\left\|g_k - A_k^T\lambda_k\right\|_2$ be bounded for all $k$ and (3.39), it follows that there exists $i_0$ such that the following property hold

$$\|\hat{g}_k\|_2 \geq \sigma_k\left\|A_k^T c_k\right\|_2 - \left\|g_k - A_k^T\lambda_k\right\|_2 \geq \frac{\epsilon}{2}\sigma_k, \quad \text{for all} \quad k \in \mathcal{L} \quad \text{and} \quad i \geq i_0. \tag{3.40}$$

Hence, by Lemma 3.3 we get

$$Pred_k \geq \frac{\bar{\beta}\epsilon\sigma_k}{4}\min\left\{\frac{\epsilon}{2\xi}, \Delta_k\right\}, \quad \text{for } k \in \mathcal{L}, i \geq i_0. \tag{3.41}$$

From this and $\{\delta_k\sigma_k\}_{\mathcal{L}} \to 0$, it follows that $Pred_k < \delta_k\sigma_k \min\left\{\Delta_k \|c_k\|_2, \|c_k\|_2^2\right\}$ only can occurs for a finite number of $k$, $k \in \mathcal{L}$. Therefore, the increase of $\sigma_k$ in $\mathcal{L}$ is due to (2.49) occur infinitely many times. Then we must have $\|\hat{g}_k\|_2 = 0$ for infinite $k \in \mathcal{L}$. This contradicts (3.40), because $\frac{\epsilon}{2}\sigma_k > 0 \ \ \forall k$. The proof is complete. ■

The following theorem presents the global convergence result of Algorithm 2.1 when $\sigma_k \to \infty$ and $\|c_k\|_2 \to 0$. Its proof is an adaptation of the proof of Theorem 3.3 in [13].

**Theorem 3.6.** *Suppose that A1-A4 hold. If* $\lim_{k\to\infty}\sigma_k = \infty$ *and* $\lim_{k\to\infty}\|c_k\|_2 = 0$, *then the sequence of iterates* $\{x_k\}$ *is not bounded away from KKT points of (1.1)-(1.2), or its Fritz-John points at which the RCPLD condition is not satisfied.*

*Proof:* Let $\bar{x}$ an accumulation point of $\{x_k\}$. Then, $c(\bar{x}) = 0$. If the RCPLD condition fails to hold at $\bar{x}$, then all the gradients of constraints at $\bar{x}$ are linearly dependent. Thus, $\bar{x}$ is a Fritz-John point. Now, assume that the RCPLD condition holds at accumulation points of $\{x_k\}$. As $\sigma_k \to \infty$ and the increase of $\sigma_k$ is due to (2.49) or (2.50), two possible cases may happen.

*Case 1*: The update (2.49) occurs at most finite number of iterations. Then, for all $k$ sufficiently large, only the update (2.50) happens at $k$th iteration. Without loss of generality, assume that (2.49) never happens. Firstly, it will be proved that for all $\epsilon > 0$, there exists $k = k(\epsilon)$, such that

$$\|c_k\|_2 < \epsilon \quad \text{and} \quad \left\|P_{\mathrm{Null}(\bar{A}_k)}\left(\bar{g}_k - \bar{A}_k^T\lambda_k\right)\right\|_2 < \epsilon. \tag{3.42}$$

Suppose, by contradiction, that (3.42) is not true. Then, due to $\lim_{k\to\infty}\|c_k\|_2 = 0$, there exists $\bar{\epsilon} > 0$ such that for all $k$ sufficiently large,

$$\|c_k\|_2 < \bar{\epsilon} \quad \text{and} \quad \left\|P_{\mathrm{Null}(\bar{A}_k)}\left(\bar{g}_k - \bar{A}_k^T\lambda_k\right)\right\|_2 \geq \bar{\epsilon}. \tag{3.43}$$

By Lemma 3.3 and by (3.43) we have

$$\begin{aligned} Pred_k &\geq \frac{\bar{\beta}}{2}\left\|P_{\bar{\mathcal{N}}_k}\left(\bar{g}_k - \bar{A}_k^T\lambda_k\right)\right\|_2 \min\left\{\frac{\left\|P_{\bar{\mathcal{N}}_k}\left(\bar{g}_k - \bar{A}_k^T\lambda_k\right)\right\|_2}{\sigma_k\xi}, \Delta_k\right\} \\ &\geq \frac{\bar{\beta}}{2}\bar{\epsilon}\min\left\{\frac{\bar{\epsilon}}{\sigma_k\xi}, \Delta_k\right\}, \end{aligned}$$

for $k$ sufficiently large. As $\sigma_k \to \infty$ when $k \to \infty$, then for $k$ sufficiently large, it follows that $\frac{1}{\sigma_k\xi}$ is small. Let $\nu = \min\left\{\frac{1}{\sigma_k\xi}, 1\right\}$ and denote $\bar{\nu} = \frac{\bar{\beta}}{2}\nu$. Thus, from last inequality and from (3.43) we obtain

$$Pred_k \geq \frac{\bar{\beta}}{2}\min\left\{\nu\Delta_k\bar{\epsilon}, \nu\bar{\epsilon}^2\right\} > \bar{\nu}\min\left\{\Delta_k\|c_k\|_2, \|c_k\|_2^2\right\}, \tag{3.44}$$

for $k$ sufficiently large. However, $\sigma_k \to \infty$ is due to (2.50), and this implies that

$$Pred_k < \delta_k\sigma_k\min\left\{\Delta_k\|c_k\|_2, \|c_k\|_2^2\right\}$$

holds for infinite $k$ with $\delta_k\sigma_k \to 0$ when $k \to \infty$. This contradicts (3.44). Therefore, for any $\epsilon > 0$, there exists $k$ such that (3.42) holds. Consequently, letting $\epsilon \to 0$ we obtain a subsequence $\mathcal{K}$ such that

$$\lim_{\substack{k\to\infty \\ k\in\mathcal{K}}}\|c_k\|_2 = 0 \quad \text{and} \quad \lim_{\substack{k\to\infty \\ k\in\mathcal{K}}}\left\|P_{\mathrm{Null}(\bar{A}_k)}\left(\bar{g}_k - \bar{A}_k^T\lambda_k\right)\right\|_2 = 0. \tag{3.45}$$

Let $x_*$ an accumulation point of $\{x_k\}_{\mathcal{K}}$. Then there exists a subset $\mathcal{K}_1 \subseteq \mathcal{K}$ such that $\{x_k\}_{\mathcal{K}_1} \to x_*$. From this and by A1 it follows that $\{\|c(x_k)\|_2\}_{\mathcal{K}_1} \to \|c(x_*)\|_2 = 0$. Assume that the RCPLD condition holds at $x_*$. It will be proved that $x_*$ is a KKT point of (1.1)-(1.2). Denote $\{\nabla c_i(x_*)\}_{i\in I}$ as the maximal set of linearly independent vectors, among all the gradient vectors of constraints at $x_*$. Then, by the definition of the RCPLD condition and by A1, there exists a neighborhood of $x_*$ such that for any point $x$ in this neighborhood, the vectors $\{\nabla c_i(x)\}_{i\in I}$ are linearly independent and

$$\text{span}\left\{\nabla c_i(x), i = 1, \ldots, m\right\} = \text{span}\left\{\nabla c_i(x), i \in I\right\}.$$

Then, there is a index $k_0$ such that for all $k_0 < k \in \mathcal{K}$, $\{\nabla c_i(x_k)\}_{i\in I}$ are linearly independent vectors and

$$\text{Range}(A_k^T) = \text{span}\left\{\nabla c_i(x_k), i = 1, \ldots, m\right\} = \text{span}\left\{\nabla c_i(x_k), i \in I\right\}. \tag{3.46}$$

Consider now a new matrix $\hat{A}_k^T$, whose columns are the vectors $\nabla c_i(x_k)$, $i \in I$ and $k_0 < k \in \mathcal{K}$. Similarly, denote $\hat{A}_*^T$ as $\hat{A}_*^T = (\nabla c_i(x_*))_{i\in I}$. Since $\hat{A}_*^T$ has full column rank and $\{x_k\}_{\mathcal{K}_1} \to x_*$, then there exists $k_1 \in \mathcal{K}_1$, $k_1 \geq k_0$ such that $\hat{A}_k^T$ has full column rank for any $k_1 \leq k \in \mathcal{K}_1$. From this and (3.46) there exists a vector $v_k$ such that

$$\bar{A}_k^T \lambda_k = Z_k^T A_k^T \lambda_k = Z_k^T \hat{A}_k^T v_k.$$

Now using the Lemma 3.2 with $\hat{A}_k^T$ we obtain

$$\left\| P_{\bar{\mathcal{N}}_k}\left(\bar{g}_k - Z_k^T \hat{A}_k^T v_k\right)\right\|_2 = \left\| P_{\mathcal{N}_k}\left(g_k - \hat{A}_k^T v_k\right)\right\|_2, \tag{3.47}$$

for $k_1 \leq k \in \mathcal{K}_1$. Hence, from (3.45) and (3.47) it follows that for any $\hat{\epsilon} > 0$

$$\begin{aligned}
&\left\| P_{\mathcal{N}_k}(g_k - \hat{A}_k^T v_k)\right\|_2 < \hat{\epsilon} \\
&\Longrightarrow \left\|(I_n - \hat{A}_k^T(\hat{A}_k^T)^\dagger)(g_k - \hat{A}_k^T v_k)\right\|_2 < \hat{\epsilon} \\
&\Longrightarrow \left\| g_k - \hat{A}_k^T v_k - \hat{A}_k^T(\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k g_k + \hat{A}_k^T(\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k\hat{A}_k^T v_k\right\|_2 < \hat{\epsilon} \\
&\Longrightarrow \left\| g_k - \hat{A}_k^T(\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k g_k\right\|_2 < \hat{\epsilon} \\
&\Longrightarrow \left\| g_k - \hat{A}_k^T y_k\right\|_2 < \hat{\epsilon}
\end{aligned} \tag{3.48}$$

for $k = k(\hat{\epsilon})$, $k_1 \leq k \in \mathcal{K}_1$ and $y_k = (\hat{A}_k\hat{A}_k^T)^{-1}\hat{A}_k g_k = (\hat{A}_k^T)^\dagger g_k$. Thus, letting $\hat{\epsilon} \to 0$, from (3.48) it follows that

$$\left\| g_k - \hat{A}_k^T y_k\right\|_2 \to 0,$$

when $k \in \mathcal{K}_1$, $k \to \infty$. By A1, we have $g(x_k) \to g(x_*)$ and $\hat{A}(x_k) \to \hat{A}(x_*)$ when $k \in \mathcal{K}_1$, $k \to \infty$. Hence, there exists $y_*$ such that $y_k \to y_*$, when $k \in \mathcal{K}_1$, $k \to \infty$. Therefore, we obtain $\left\| g_* - \hat{A}_*^T y_* \right\|_2 = 0 \Rightarrow g_* - \hat{A}_*^T y_* = 0$. This shows that $g_* \in \text{span}\{\nabla c_i(x_*), i \in I\}$, which implies $g_* \in \text{Range}(A_*^T)$. Therefore, there exists $\mu_*$ such that $g_* = A_*^T \mu_*$ and as $c_* = 0$ we conclude that $x_*$ is a KKT point of (1.1)-(1.2).

*Case 2*: The update (2.49) occurs infinitely many times. In this case, by the update (2.49), it follows that there exist subsequences $\{x_k\}_\mathcal{K}$ and $\{\bar{\sigma}_k\}_\mathcal{K}$ such that

$$g_k - A_k^T \lambda_k + \bar{\sigma}_k A_k^T c_k = 0, \qquad k \in \mathcal{K}. \tag{3.49}$$

Suppose that $x_*$ is an accumulation point of $\{x_k\}_\mathcal{K}$ and $\{x_k\}_{\mathcal{K}_1} \to x_*$ where $\mathcal{K}_1 \subseteq \mathcal{K}$. If the RCPLD condition holds at $x = x_*$, for any $k \in \mathcal{K}$ sufficiently large, similar to the analysis in *Case 1*, there exists $\hat{A}_k$ with full row rank such that

$$\text{Range}(A_k^T) = \text{Range}(\hat{A}_k^T).$$

As (3.49) indicates that $g_k \in \text{Range}(A_k^T)$, consequently $g_k \in \text{Range}(\hat{A}_k^T)$. Therefore, $g_* \in \text{Range}(\hat{A}_*^T) \Rightarrow g_* \in \text{Range}(A_*^T)$ by the fact $\{x_k\}_{\mathcal{K}_1} \to x_*$. From this and as $c_* = 0$, we conclude that $x_*$ is a KKT point of (1.1)-(1.2). ■

## 3.2 Case II: Sequence of penalty parameters is bounded

When all the penalty parameters $\sigma_k$ are bounded this means that $\sigma_k$ keeps unchanged for all large $k$, equivalently, neither (2.49) or (2.50) happens. Without loss of generality, assume that

$$\sigma_k = \sigma \quad \text{for all } k. \tag{3.50}$$

Consequently, assume also that $\delta_k = \delta$ for all $k$.

The first result guarantees that all accumulation points of the sequence $\{x_k\}$ are feasible points if $\{\sigma_k\}$ is bounded. Its proof is due to Wang and Yuan [13]. We give it here for completeness.

**Lemma 3.7.** *Suppose that A1-A3 hold. If the sequence $\{\sigma_k\}$ is bounded, then*

$$\lim_{k\to\infty} \|c_k\|_2 = 0. \tag{3.51}$$

*Proof:* As (2.50) never happens, we have that

$$Pred_k \geq \delta\sigma \min\left\{\Delta_k \|c_k\|_2, \|c_k\|_2^2\right\}. \tag{3.52}$$

By the update of $\lambda_k$ in Step 4 of Algorithm 2.1 and by A2 if follows that

$$\sum_{k=1}^{\infty} \left(-\lambda_k^T c_k + \lambda_k^T c_{k+1}\right) \leq 2 \left\|\lambda_{\max}\right\|_2 \left(\left\|c_{\max}\right\|_2 + \frac{R_1}{1-\beta}\right) < \infty$$

Hence, similar to the analysis in Lemma 3.4, the sum of all $Ared_k$ is bounded

$$\sum_{k=1}^{\infty} Ared_k = \sum_{k=1}^{\infty} (f_k - f_{k+1}) + \sum_{k=1}^{\infty} \left(-\lambda_k^T c_k + \lambda_k^T c_{k+1}\right) + \frac{\sigma}{2} \sum_{k=1}^{\infty} \left(\|c_k\|_2^2 - \|c_{k+1}\|_2^2\right)$$

$$\leq \bar{M} \tag{3.53}$$

As $Ared_k = L(x_k, \lambda_k, \sigma_k) - L(x_{k+1}, \lambda_k, \sigma_k) \geq 0 \quad \forall\, k$, and the sum above is bounded, then the series $\sum_{k=1}^{\infty} Ared_k$ is convergent and, therefore, $Ared_k \to 0$ when $k \to \infty$.

In order to prove (3.51), first it will be proved that

$$\liminf_{k\to\infty} \left\|c(x_k)\right\|_2 = 0. \tag{3.54}$$

By contradiction, assume that (3.54) is not true. Then, there exist $k_0 \in \mathbb{N}$ and a constant $\tau > 0$ such that $\|c_k\|_2 \geq \tau$, for all $k > k_0$. In this case, from (3.52) we have

$$Pred_k \geq \delta\sigma \min\left\{\Delta_k \tau, \tau^2\right\}. \tag{3.55}$$

Denote $\bar{S}$ as the set of all the indexes corresponding to successful iterations, namely,

$$\bar{S} = \{k \in \mathbb{N}; \rho_k \geq \eta\}.$$

From the fact that $Ared_k \to 0$ for $k$ sufficiently large, and from (3.55) we have that

$$Ared_k = \rho_k Pred_k \geq \eta Pred_k \geq \eta\delta\sigma \min\left\{\Delta_k \tau, \tau^2\right\} \quad \text{for } k \in \bar{S}.$$

Therefore, for $k$ sufficiently large, $k \in \bar{S}$, we conclude that $\{\Delta_k\}_{k\in\bar{S}} \to 0$. Then, for the other $k$ that are not in $\bar{S}$, it follows by the update rule of the trust region radio that

$$\Delta_k \to 0 \tag{3.56}$$

for all $k$ sufficiently large. This implies that $s_k \to 0$. From this, and by A1-A2, there exists a constant $M > 0$ such that

$$\begin{aligned}|Ared_k - Pred_k| &= |L(x_k, \lambda_k; \sigma_k) - L(x_k + s_k, \lambda_k; \sigma_k) - q_k(0) + q(s_k)| \\ &= |f(x_k) - f(x_k + s_k) - \lambda_k^T c(x_k) + \lambda_k^T c(x_k + s_k) - \frac{\sigma}{2} \|c(x_k + s_k)\|_2^2 \\ &\quad - g(x_k)^T s_k + \lambda_k^T A_k s_k - \frac{1}{2} s_k^T B_k s_k - \sigma c_k^T A_k s_k + \frac{\sigma}{2} \|A_k s_k\|_2^2 | \\ &\leq M \|s_k\|_2^2 \\ &\leq M \Delta_k^2 \end{aligned} \tag{3.57}$$

and then

$$|\rho_k - 1| = \frac{|Ared_k - Pred_k|}{Pred_k} \leq \frac{M \Delta_k^2}{\delta \sigma \min \{\Delta_k \tau, \tau^2\}} \rightarrow 0.$$

This implies that $\rho_k \rightarrow 1$, and then we have that $\Delta_{k+1} \geq \Delta_k$ for all $k$ sufficiently large, which contradicts (3.56). Therefore, (3.54) is true.

Now, we will prove (3.51). Again, by contradiction, suppose that (3.51) is not true. Then there exists an infinite set of indexes $\{m_i\} \subset \bar{S}$ and a constant $\nu > 0$ such that

$$\|c_{m_i}\|_2 \geq 2\nu. \tag{3.58}$$

Further, (3.54) ensures the existence of a subsequence $\{n_i\} \subset \bar{S}$ where $n_i$ is the first index bigger than $m_i$ such that

$$\|c_k\|_2 \geq \nu \quad \text{for } m_i \leq k < n_i \quad \text{and} \quad \|c_{n_i}\|_2 < \nu. \tag{3.59}$$

Define the set $\mathcal{K} = \bigcup_i \{k \in \bar{S}; m_i \leq k < n_i\}$. From (3.52), (3.58) we have that for $k \in \mathcal{K}$

$$Ared_k \geq \eta \delta \sigma \min \{\Delta_k \nu, \nu^2\}. \tag{3.60}$$

Let $\bar{\nu} = \min \{\nu, \nu^2\}$. Then, we obtain

$$Ared_k \geq \eta \delta \sigma \min \{\Delta_k \bar{\nu}, \bar{\nu}\} = \eta \delta \sigma \bar{\nu} \min \{\Delta_k, 1\}$$

and setting $\zeta = \eta \delta \sigma \bar{\nu} > 0$ it follows from (3.53) that $Ared_k \leq \zeta$ for $k$ sufficiently large. From this, it follows that

$$\zeta > Ared_k \geq \zeta \min \{\Delta_k, 1\} \Rightarrow \min \{\Delta_k, 1\} < 1 \Rightarrow \min \{\Delta_k, 1\} = \Delta_k$$

therefore

$$Ared_k \geq \zeta \Delta_k, \tag{3.61}$$

for $k \in \mathcal{K}$, $k$ sufficiently large. From (3.61) and $Ared_k \to 0$ for $k$ sufficiently large, we have that $\lim_{\substack{k\to\infty \\ k\in\mathcal{K}}} \Delta_k = 0$. Then, for any $\epsilon > 0$ there exists $i$ sufficiently large, (see Theorem 17 in [6]) such that

$$\begin{aligned}
\|x_{m_i} - x_{n_i}\|_2 &= \left\|x_{m_i} - x_{m_{i+1}} + x_{m_{i+1}} - \ldots - x_{n_{i-1}} + x_{n_{i-1}} - x_{n_i}\right\|_2 \\
&\leq \sum_{\substack{k=m_i \\ k\in\bar{S}}}^{n_i-1} \|x_k - x_{k+1}\|_2 \leq \sum_{k=m_i}^{n_i-1} \|s_k\|_2 \leq \sum_{k=m_i}^{n_i-1} \|\Delta_k\|_2 \\
&\leq \frac{1}{\zeta} \sum_{\substack{k=m_i \\ k\in\bar{S}}}^{n_i-1} Ared_k \leq \frac{1}{\zeta}\epsilon\zeta = \epsilon.
\end{aligned} \tag{3.62}$$

Hence, letting $\epsilon \to 0$ it follows that $\|x_{m_i} - x_{n_i}\|_2 \to 0$ when $i \to \infty$. From this and by A1 we have that $\|c_{m_i} - c_{n_i}\|_2 \to 0$ as $i \to \infty$. Note that

$$\|c_{m_i}\|_2 = \|c_{m_i} - c_{n_i} + c_{n_i}\|_2 \leq \|c_{m_i} - c_{n_i}\|_2 + \|c_{n_i}\|_2 \leq \frac{\nu}{2} + \nu < 2\nu \tag{3.63}$$

which contradicts (3.58). The proof is complete. ■

Finally, the next result establishes the global convergence of Algorithm 2.1 when the penalty parameters are bounded. The proof is an adaptation of Theorem 3.5 from [13].

**Theorem 3.8.** *Suppose that A1-A4 hold. If the sequence $\{\sigma_k\}$ is bounded, then the sequence $\{x_k\}$ generate by Algorithm 2.1 is not bounded away from KKT points of (1.1)-(1.2).*

*Proof:* By Lemma 3.3, we have that

$$Pred_k \geq \frac{\bar{\beta}}{2} \|\hat{g}_k\|_2 \min\left\{\frac{\|\hat{g}_k\|_2}{\sigma\xi}, \Delta_k\right\}. \tag{3.64}$$

To show the enunciated result, we will show that

$$\liminf_{k\to\infty} \|\hat{g}_k\|_2 = 0. \tag{3.65}$$

Suppose, by contradiction, that (3.65) is not true. Then, there exist a constant $\epsilon > 0$ and a index $k_0$ such that

$$\|\hat{g}_k\|_2 \geq \epsilon \quad \text{for } \ \forall\, k \geq k_0. \tag{3.66}$$

From (3.64) and (3.66), we have that, for $k \geq k_0$,

$$Pred_k \geq \frac{\bar{\beta}}{2}\epsilon \min\left\{\frac{\epsilon}{\sigma\xi}, \Delta_k\right\}. \tag{3.67}$$

Consider the set $\bar{S}$ as in previous Lemma 3.7. For $k \in \bar{S}$, $k \geq k_0$ we have that

$$Ared_k \geq \eta Pred_k \geq \eta\frac{\bar{\beta}}{2}\epsilon \min\left\{\frac{\epsilon}{\sigma\xi}, \Delta_k\right\}. \tag{3.68}$$

As seen in Lemma 3.7, $Ared_k \to 0$. Then, it implies that $\{\Delta_k\}_{\bar{S}} \to 0$ and, consequently, $\Delta_k \to 0$ as $k \to \infty$. This implies that $\|s_k\|_2 \to 0$ as $k \to \infty$. Thus, by A1-A2, there exists $M > 0$ such that

$$|\rho_k - 1| = \frac{|Ared_k - Pred_k|}{Pred_k} \leq \frac{M\Delta_k^2}{\frac{\bar{\beta}}{2}\epsilon \min\left\{\frac{\epsilon}{\sigma\xi}, \Delta_k\right\}} \to 0 \tag{3.69}$$

and, then, $\rho_k \to 1$ when $k \to \infty$. This implies that $\Delta_{k+1} \geq \Delta_k$ for all $k$ sufficiently large, contradicting the fact that $\Delta_k \to 0$ when $k \to \infty$. Therefore (3.65) is true.

Now, note that

$$\left\|g_k - A_k^T\lambda_k\right\|_2 = \left\|g_k - A_k^T\lambda_k + \sigma A_k^T c_k - \sigma A_k^T c_k\right\|_2 \leq \|\hat{g}_k\|_2 + \sigma\left\|A_k^T c_k\right\|_2. \tag{3.70}$$

Let $\epsilon_1 > 0$ arbitrary. Then, combining the fact that $\|A_k\|_2 \leq \kappa_1$ for all $k$, with (3.65) and Lemma 3.7, it follows that there exist indices $k_1$ and $k_2$ such that

$$\|\hat{g}_k\|_2 \leq \frac{\epsilon_1}{2} \quad \text{for} \quad k \geq k_1 \quad \text{and} \quad \|c_k\|_2 \leq \frac{\epsilon_1}{2\sigma\kappa_1} \quad \text{for} \quad k \geq k_2. \tag{3.71}$$

Now, from (3.70) and (3.71) we obtain

$$\left\|g_k - A_k^T\lambda_k\right\|_2 \leq \frac{\epsilon_1}{2} + \sigma\frac{\epsilon_1}{2\sigma} = \epsilon_1 \quad \text{for} \quad k > \max\{k_1, k_2\}. \tag{3.72}$$

Therefore,

$$\liminf_{k\to\infty}\left\|g_k - A_k^T\lambda_k\right\|_2 = 0.$$

By Lemma 3.7, we have $\|c_k\|_2 \to 0$ as $k \to \infty$. Thus, there exists a subsequence of $\{x_k\}$, say $\{x_{k_j}\}$, with $x_{k_j} \to x_*$ and $\left\|g_{k_j} - A_{k_j}^T\lambda_{k_j}\right\|_2 \to 0$ as $j \to \infty$. Consequently, $g(x_*) = A(x_*)^T\lambda_*$ and $c(x_*) = 0$. Therefore, $\{x_k\}$ has an accumulation point, $x_*$, which is a KKT point of (1.1)-(1.2). ■

# Chapter 4

# Illustrative Numerical Experiments

To investigate the advantages and limitations of our subspace algorithm we have tested MATLAB implementatios of Algorithms 1.1 and 2.1, which are referred as “ALTR” and “SALTR”, respectively. In both implementations we consider the parameters $\beta = 0.99$, $\theta = 10$, $\eta = 10^{-8}$, $\eta_1 = 0.1$, $\lambda_1 = \mathbb{1}$, $\sigma_1 = 10$, $\Delta_1 = 1$ and $\delta_1 = 0.1$. In ALTR, the initial matrix $B_1$ is chosen as the $n$-dimensional identity matrix and $B_k$ is updated using the BFGS formula. In SALTR, the initial matrix $\bar{B}_1$ is also chosen as the identity matrix, however its dimension is equal to the dimension of the first subspace (i.e., $G_1$) while $\bar{B}_k$ is updated by the BFGS formula in the subspace. For both codes, the execution is interrupted when any of the conditions below is satisfied:

$$\max\left\{\|c_k\|_2, \left\|P_{\mathrm{Null}(A_k)}(g_k)\right\|_2\right\} \leq 10^{-5}, \tag{4.1}$$

$$\max\{\|c_k\|_2, \|s_k\|_2\} \leq 10^{-5}, \tag{4.2}$$

$$\|s_k\|_2 \leq 10^{-15}, \tag{4.3}$$

$$k \geq 10,000. \tag{4.4}$$

The tests were performed with MATLAB 8.5.0 (R2015a), on a PC with a 2.50 GHz Intel(R) Core(TM) i5-3210 microprocessor, and 6GB of memory.

## 4.1 Advantages of our subspace algorithm

First, we applied ALTR and SALTR to a set of 10 nonlinear least-square problems with a single equality constraint, i.e., problems of the form:

$$\min_{x \in \mathbb{R}^n} \quad \sum_{i=1}^{p} [f_i(x)]^2 , \tag{4.5}$$

$$\text{s. t.} \quad c(x) = 0, \tag{4.6}$$

where $f, c : \mathbb{R}^n \rightarrow \mathbb{R}$. Specifically, each problem was formed with functions from [8]. Problems and the corresponding initial points are described in the Appendix. The results for $n = 1,000$ are given in Table 4.1, where "IT" represents the number of iterations, "TIME" represents the CPU time (in seconds), "$f^*$" represents the final objective value, "$c^*$" represents the infinity norm of the final constraint violation, and "TOTAL" gives the sum of the values in the corresponding column.

| | **ALTR** | | | | **SALTR** | | | |
|---|---|---|---|---|---|---|---|---|
| **PROBLEM** | **IT** | **TIME** | $f^*$ | $c^*$ | **IT** | **TIME** | $f^*$ | $c^*$ |
| 1 | 6 | 5.4 | 2.1101E+03 | 1.6420E-01 | 8 | 0.1 | 2.1100E+03 | 1.6930E-01 |
| 2 | 2434 | 1128.1 | 8.3000E-04 | 2.1000E-07 | 2509 | 161.4 | 7.8000E-05 | 9.0000E-08 |
| 3 | 2451 | 332.2 | 9.9000E+05 | 4.0000E-06 | 2643 | 34.6 | 9.9000E+05 | 3.7000E-06 |
| 4 | 888 | 135.4 | 6.6000E+22 | 1.9000E-07 | 1015 | 1.1 | 6.6000E+22 | 8.3000E-09 |
| 5 | 4405 | 1660.4 | 4.3097E+15 | 1.4954E-07 | 8244 | 95.6 | 6.9448E+13 | 3.7738E-10 |
| 6 | 34 | 8.3 | 1.5443E+04 | 7.6932E-06 | 42 | 0.13 | 1.5428E+04 | 7.0606E-06 |
| 7 | 9 | 10.5 | 5.7400E-02 | 1.9000E-08 | 9 | 0.1 | 3.4700E-02 | 1.1000E-08 |
| 8 | 15 | 7.04 | 9.8924E+02 | 2.1020E-01 | 19 | 0.14 | 9.8925E+02 | 2.1000E-01 |
| 9 | 17 | 46.4 | 1.2643E-07 | 7.0079E-10 | 15 | 0.96 | 1.2617E-09 | 2.3903E-09 |
| 10 | 32 | 6.2 | 9.3775E+02 | 1.5000E-08 | 13 | 0.3 | 1.1000E+03 | 2.1000E-10 |
| **TOTAL** | - | 3339.94 | - | - | - | 294.43 | - | - |

Table 4.1: Numerical Results for Problems 1–10

From Table 4.1 we can see that ALTR and SALTR stopped with comparable approximate solutions. However SALTR is significantly more efficient than ALTR in terms of CPU time. In particular, SALTR reduced the total time (taken to solve all problems) in 91.2% with respect to ALTR.

In addition, we have tested Problem 9 for $n = 1,000,000$. In this case, it was not possible to run code ALTR, since the initial matrix $B_1$ exceeded the maximum array size allowed by MATLAB. In contrast, despite the large dimension of the problem, our subspace code was able to return an approximate solution in 306.9 seconds (22 iterations), with $f^* = 3.8956E-10$ and $c^* = 5.7799E-09$. This was possible because the dimension of the subspace $G_k$ remained much smaller than $n$ during the execution of SALTR, as we can see in Figure 4.1.

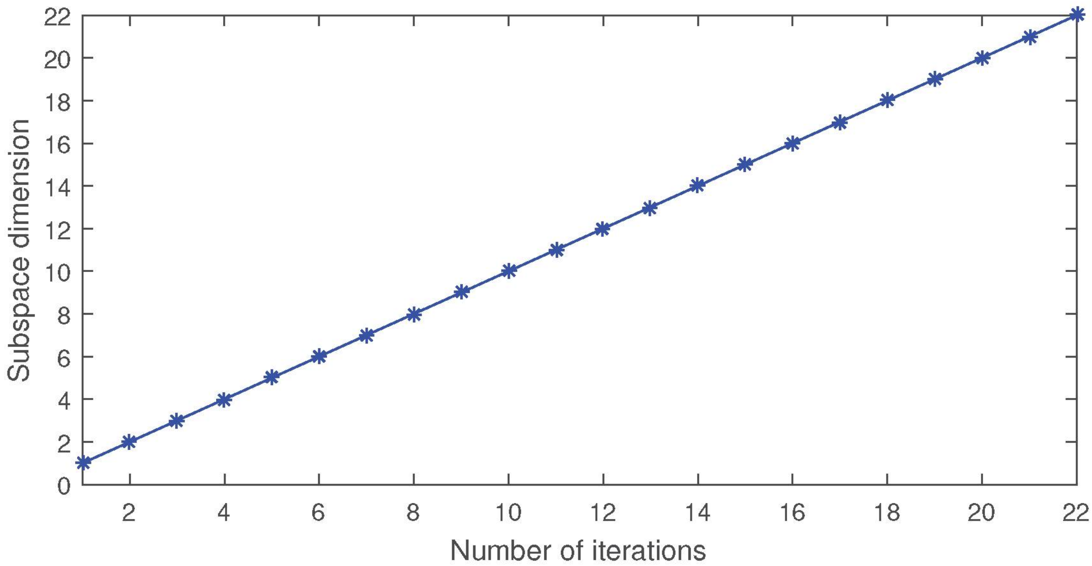



Figure 4.1: Growth of the subspace dimension with the number of iterations

These results show that our subspace algorithm can provide a significant reduction in the computational time in comparison to its "full space" counterpart, when the number of constraints is much lower than the number of variables.

## 4.2 Limitations of our subspace algorithm

In order to evaluate the influence of the number of constraints ($m$) on the performance of SALTR, we applied ALTR and SALTR to the following class of problems:

$$\min_{x\in\mathbb{R}^n} \quad \sum_{i=1}^{n} [f_i(x)]^2\,, \tag{4.7}$$

$$\text{s. t.} \quad c_j(x) = 0,\ j = 1,\dots,m \tag{4.8}$$

where, for $i = 1,\dots,n$,

$$f_i(x) = \begin{cases} 10(x_{i+1} - x_i^2), & \text{if } i \text{ is odd,} \\ 1 - x_{i-1}, & \text{otherwise,} \end{cases} \tag{4.9}$$

for $j = 1,\dots,m$,

$$c_j(x) = n - \sum_{i=1}^{n} \cos(x_i) + j(1 - \cos(x_j)) - \sin(x_j), \tag{4.10}$$

and the initial point is $x_1 = (\xi_i)$, with $\xi_{2i-1} = -1.2$ and $\xi_{2i} = 1$. The results for $n = 1,000$ are shown in Table 4.2.

| | ALTR | | | | SALTR | | | |
|---|---|---|---|---|---|---|---|---|
| **m** | **IT** | **TIME** | $f^*$ | $c^*$ | **IT** | **TIME** | $f^*$ | $c^*$ |
| 1 | 21 | 7.9 | 4.6896E+02 | 3.2000E-02 | 22 | 0.06 | 4.6922E+02 | 3.4000E-02 |
| 100 | 60 | 19.9 | 4.9685E+02 | 1.5900E-02 | 41 | 1.23 | 4.9639E+02 | 1.3500E-02 |
| 250 | 2193 | 349.3 | 4.9745E+02 | 6.7503E-04 | 2206 | 226.1 | 4.9755E+02 | 1.1811E-04 |

Table 4.2: Numerical Results for problems (4.9)-(4.10)

As expected, the advantage of SALTR over ALTR decreases as we increase the number of constraints. This observation is highlighted in Figure 4.2, which shows the percentage reduction in CPU time that SALTR promotes (with respect to ALTR) as a function of the number of constraints.

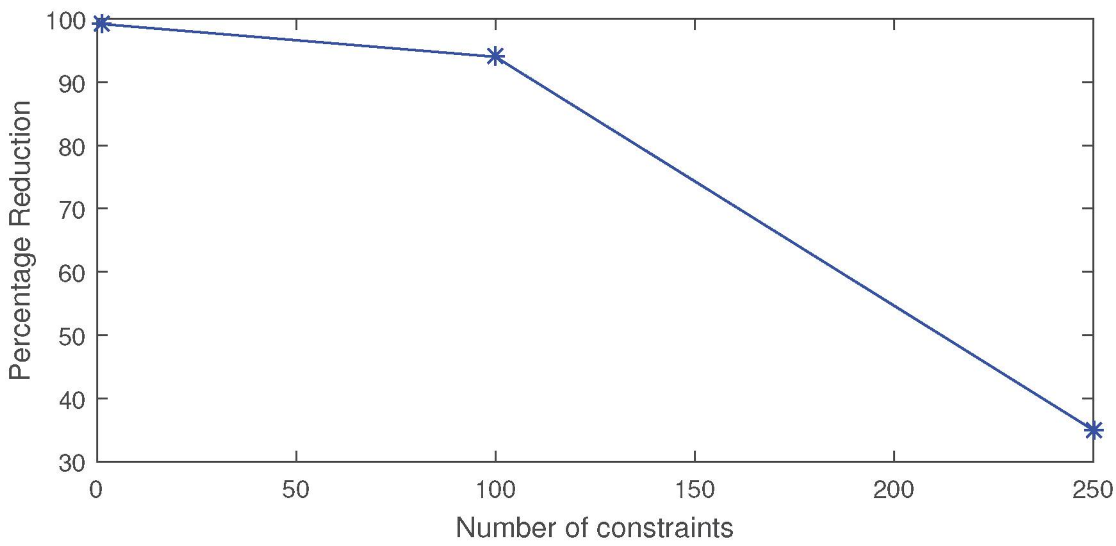



Figure 4.2: Percentage reduction in CPU time for problems (4.9)-(4.10)

In summary, as anticipated by our theoretical analysis, it is safe to recommend the use of SALTR instead of ALTR *only when* the number of equality constraints is much lower than the number of variables.

# Conclusion

In this dissertation, a subspace version of the Augmented Lagrangian-Trust Region method (ALTR) in [13] is proposed for large-scale equality constrained optimization problems. The subspace scheme is based on the finding that, at the $k$th iteration of ALTR, any solution $s_k$ of the corresponding trust-region subproblem belongs to the subspace

$$G_k = \text{span}\left(\cup_{i=1}^{k} \{\nabla c_1(x_i), \ldots, \nabla c_m(x_i), g_i\}\right),$$

as long as the approximate Hessians are updated by suitable quasi-Newton formulas. Thus, in the subspace variant of ALTR (called SALTR), the computation of $s_k$ is restricted to the subspace $G_k$. By adapting the analysis from [13], the global convergence of the subspace ALTR is established assuming inexact solution of the subproblems and considering a subspace more general than $G_k$.

In line with the theory, illustrative numerical experiments show that the subspace algorithm can outperform its "full space" counterpart on problems in which the number of constraints is much lower than the number of variables.

Future research include the development of a strategy to control the dimension of the subspaces (similar to [5]), and also the extension of this subspace approach to problems with inequality constraints [14].

# References


[1] J. W. Daniel, W. B. Gragg, L. Kaufman and G. W. Stewart. *Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization.* Mathematics of Computation, 30: 772-795, 1976.

[2] P. E. Gill and M. W. Leonard. *Reduced-Hessian quasi-Newton methods for unconstrained optimization.* SIAM Journal on Optimization, 12: 209-237, 2001.

[3] G. H. Golub and C. F. Van Loan. *Matrix Computations, 3rd edition.* The Johns Hopkins University Press, 1996.

[4] G. N. Grapiglia, J. Yuan and Y. Yuan. *A subspace version of the Powell-Yuan algorithm for equality-constrained optimization.* Journal of the Operations Research Society of China, 1: 425-451, 2013.

[5] L. Gong. *A trust region subspace method for large-scale unconstrained optimization.* Asia-Pacific Journal of Operational Research, 29: 1-16, 2012.

[6] E. L. Lima. *Curso de Análise, Vol. 1.* IMPA, 2016.

[7] J. J. Moré and D.C. Sorensen. *Computing a trust region step.* SIAM Journal on Scientific and Statistical Computing, 4: 553-572, 1983.

[8] J. J. Moré, B. S. Garbow and K. E. Hillstrom. *Testing Unconstrained Optimization Software.* ACM Transactions on Mathematical Software, 7: 17-41, 1981.

[9] A. A. Ribeiro and E. W. Karas, *Otimização Contínua: aspectos teóricos e computacionais.* Cengage Learning, 2013.

[10] D. Siegel. *Implementing and Modifying Broyden class updates for large scale optimization.* Report DAMPT 1992/NA12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England, 1992.

[11] W. Sun and Y. Yuan. *Optimization Theory and Methods: Nonlinear Programming.* Springer, 2006.

[12] Z. H. Wang and Y. Yuan. *A subspace implementation of quasi-Newton trust region methods for unconstrained optimization.* Numerische Mathematik, 104: 241-269, 2006.

[13] X. Wang and Y. Yuan. *An augmented Lagrangian trust region method for equality constrained optimization.* Optimization Methods and Software, 30:3, 559-582, 2015.

[14] X. Wang and H. Zhang. *An augmented Lagrangian affine scaling method for nonlinear Programming.* Optimization Methods and Software, 30: 934-964, 2015.

# Appendix A

# Test Problems

Below, we describe in detail the test problems considered in Chapter 4.

**Problem 1**

**(i)** $n$ even, $p = n$

**(ii)** $f_{2i-1}(x) = 10(x_{2i} - x_{2i-1}^2)$
$f_{2i}(x) = 1 - x_{2i-1}$

**(iii)** $c(x) = \left(\sum_{j=1}^{n} x_j^2\right) - n$

**(iv)** $x_1 = (\xi_j)$ where $\xi_{2j-1} = -1.2,\ \xi_{2j} = 1$

**Problem 2**

**(i)** $n$ multiple of 4, $p = n$

**(ii)** $f_{4i-3}(x) = x_{4i-3} + 10x_{4i-2}$
$f_{4i-2}(x) = 5^{1/2}(x_{4i-1} - x_{4i})$
$f_{4i-1}(x) = (x_{4i-2} - 2x_{4i-1})^2$
$f_{4i}(x) = 10^{1/2}(x_{4i-3} - x_{4i})^2$

**(iii)** $c(x) = \left(\sum_{j=1}^{n}(n - j + 1)x_j^2\right) - 1$

**(iv)** $x_1 = (\xi_j)$ where $\xi_{4j-3} = 3$, $\xi_{4j-2} = -1$, $\xi_{4j-1} = 0$, $\xi_{4j} = 1$

**Problem 3**

**(i)** $p = n + 1$

**(ii)** $f_i(x) = a^{1/2}(x_i - 1),$ for $1 \le i \le n$ where $a = 10^{-5}$

$$f_{n+1}(x) = \left(\sum_{j=1}^{n} x_j^2\right) - 1/4$$

**(iii)** $c(x) = \sum_{j=1}^{n} j(x_j - 1)^2$

**(iv)** $x_1 = (\xi_j)$ where $\xi_j = j$

**Problem 4**

**(i)** $p = n + 2$

**(ii)** $f_i(x) = x_i - 1,$ for $1 \le i \le n$

$$f_{n+1}(x) = \sum_{j=1}^{n} j(x_j - 1)$$

$$f_{n+2}(x) = \left(\sum_{j=1}^{n} j(x_j - 1)\right)^2$$

**(iii)** $c(x) = n - \sum_{j=1}^{n} \cos(x_j) + n(1 - \cos(x_n)) - \sin(x_n)$

**(iv)** $x_1 = (\xi_j)$ where $\xi_j = 1 - (j/n)$

**Problem 5**

**(i)** $p = n + 1$

**(ii)** $f_i(x) = a^{1/2}(x_i - 1),$ for $1 \le i \le n$ where $a = 10^{-5}$

$$f_{n+1}(x) = \left(\sum_{j=1}^{n} x_j^2\right) - 1/4$$

**(iii)** $c(x) = n - \sum_{j=1}^{n} \cos(x_j) + n(1 - \cos(x_n)) - \sin(x_n)$

**(iv)** $x_1 = (\xi_j)$ where $\xi_j = j$

**Problem 6**

**(i)** $n$ multiple of 4, $p = n$

**(ii)** $f_{4i-3}(x) = x_{4i-3} + 10x_{4i-2}$

$f_{4i-2}(x) = 5^{1/2}(x_{4i-1} - x_{4i})$

$f_{4i-1}(x) = (x_{4i-2} - 2x_{4i-1})^2$

$f_{4i}(x) = 10^{1/2}(x_{4i-3} - x_{4i})^2$

**(iii)** $c(x) = x_n(2 + 5x_n^2) + 1 - \sum_{j=1}^{n} x_j(1 + x_j)$

**(iv)** $x_1 = (\xi_j)$ where $\xi_{4j-3} = 3$, $\xi_{4j-2} = -1$, $\xi_{4j-1} = 0$, $\xi_{4j} = 1$

**Problem 7**

**(i)** $p = n$

**(ii)** $f_i(x) = 2x_i - x_{i-1} - x_{i+1} + \dfrac{h^2(x_i + t_i + 1)^3}{2}$

where $h = \dfrac{1}{n+1}$, $t_i = ih$ and $x_0 = x_{n+1} = 0$

**(iii)** $c(x) = (n-1)\left(\sum_{j=2}^{n-1} jx_j\right) - 1$

**(iv)** $x_1 = (\xi_j)$ where $\xi_j = t_j(t_j - 1)$

**Problem 8**

**(i)** $p = n$

**(ii)** $f_i(x) = (3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1$

where $x_0 = x_{n+1} = 0$

**(iii)** $c(x) = x_n(2 + 5x_n^2) + 1 - \sum_{j=1}^{n} x_j(1 + x_j)$

**(iv)** $x_1 = (-1, \ldots, -1)$

**Problem 9**

**(i)** $p = n$

**(ii)** $f_i(x) = 2x_i - x_{i-1} - x_{i+1} + \dfrac{h^2(x_i + t_i + 1)^3}{2}$

where $h = \dfrac{1}{n+1}$, $t_i = ih$ and $x_0 = x_{n+1} = 0$

**(iii)** $c(x) = \sum_{j=1}^{n} x_j^2 - 1$

**(iv)** $x_1 = (\xi_j)$ where $\xi_j = t_j(t_j - 1)$

**Problem 10**

**(i)** $p \geq n$

**(ii)** $f_i(x) = x_i - \frac{2}{p}\left(\sum_{j=1}^{n} x_j\right) - 1$

**(iii)** $c(x) = \sum_{j=1}^{n} x_j^2 - 1$

**(iv)** $x_1 = (1, \ldots, 1)$